Beyond Self-Regulated Learning: Integrating Approaches to Self-Regulation in Education

Dissertation
der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Franz Dietmar Wortha
aus Bad Muskau

Tübingen
2021

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	21.05.2021
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter:	Prof. Dr. Peter Gerjets
2. Berichterstatter:	Prof. Dr. Katharina Scheiter

Dedication

To Elfriede Leske who could not witness of this academic adventure.

Acknowledgements

First, I want to express my gratitude to my supervisors Prof. Peter Gerjets and Prof. Enkelejda Kasneci for their guidance and support. Thank you for your valuable feedback and the inspiring discussions and projects that developed during my PhD project.

Second, I like to thank Prof. Roger Azevedo, his team, and Prof. Michelle Taub at the University of Central Florida for our close collaboration. My visit at their laboratory was one of the highlights of my PhD project and I truly enjoy working with you.

Special thanks go to Dr. Birgit Brucker for her tireless support throughout the final stages of my PhD project. Further, I wish to thank Lydia Kastner for the mutual support she provided throughout our PhD projects.

Furthermore, I want to thank the LEAD Graduate School and Research Network and the Hector Research Institute for Education Sciences and Psychology for providing the interdisciplinary context that enabled a project at such a broad scope. Particular thanks to Prof. Ulrich Trautwein and Prof. Benjamin Nagengast for providing PhD students with opportunities to collaborate with leading experts in the field. I also would like to extend my thanks to Katharina Scheiter for agreeing to review this dissertation. Further, I'd like to thank Heiko Holz, Christian Mychajliw, and Kristina Davidoswki for the enjoyable work on our joint project and the friendship we share. Additional thanks to Jakob Schwerter, who I enjoyed collaborating with in our joint research project.

Lastly, my biggest thanks go to my wife Silke, who supported me through all ups and downs of my PhD project. Without your support I would not be the person I am today. For that I am forever grateful.

Abstract

Self-regulated learning (SRL) has become one of the most important theoretical concepts in educational research. In light of contemporary educational challenges, including the widespread use of information technology in educational settings, the growing focus on enabling students to become lifelong learners, and the increased emphasis on learner-controlled learning activities, SRL further shows a significant practical importance. The ability to effectively regulate learning processes is a key skill for learners to meet the aforementioned challenges. Typically, SRL is referred to as the regulation and control of cognitive, metacognitive, motivational, as well as affective states and processes in service of learning goals. Following this definition, a broad body of literature investigating SRL from different theoretical backgrounds and perspectives has shown that SRL is key factor for students' academic success throughout all stages of education. However, the diversity in approaches to investigate SRL has also led to lack of clarity what SRL is and how it can be most effectively fostered. This issues becomes even more apparent when SRL is investigated in the context of other, more general research traditions on self-regulation (SR). The present dissertation addressed this research issue by integrating four areas of research on (SR) in education. These were derived from the mechanisms through which self-regulatory variables affect learning and include learning activities (e.g., cognitive and metacognitive strategies), driving forces (e.g., motivation and affect), personal dispositions (e.g., personality), and limited resources (e.g., working memory and executive functions). Specifically, based on research that has strongly linked each of these areas of research to learning and academic achievement, an integrative framework that situates SRL as part of SR in education has been proposed. To test this framework, the present dissertation tested the predictive value of key constructs representing all areas of proposed framework across different contexts (e.g., learning in school and laboratory learning task). Through this approach, this dissertation is the first study that empirically integrated the aforementioned research traditions on self-regulation in education.

Study I aimed at identifying the best predictors of learning in school and laboratory learning tasks from a comprehensive set of self-regulatory constructs that reflect the four areas of research on self-regulation proposed in the framework (i.e.,

learning activities, driving forces, personal dispositions, and limited resources). Specifically, robust machine learning predictions were used to predict performance in school and laboratory learning task across five academic domains (i.e., math, physics, biology, art, and history). Results showed that predictors from all areas of the proposed framework are required to optimally predict learning in both settings. However, the specific variables that optimally predicted learning in school and laboratory learning tasks varied. While measures of driving forces (i.e., motivation) and limited resources (i.e., working memory capacity) predicted learning in both settings, predictors representing learning activities (e.g., effort-related vs rehearsal strategies) and personality (e.g., openness) only showed predictive value for one of the outcomes.

Study II investigated if and how self-regulatory requirements in a computer-based learning task differed depending on the way participants interacted with the learning environment. In detail, participants used either mouse-based or touch-based interaction to work with the learning materials. Robust machine learning models predicting learning outcomes in both conditions were developed. Specifically, these models used measures that represent the four core areas of the proposed framework similar to Study I. Results showed that self-regulatory requirements were higher when learning with tablets. Specifically, beyond the predictive value of prior knowledge, learning on tablet was determined by critical evaluation (learning activity), motivational cost (driving force), openness (personal disposition), and switching (limited resource). Differences in performance using mouse-based interactions on the other hand were only related to control measures (reading comprehension and prior knowledge) but not related to self-regulatory constructs.

Study III extended the scope of the first two studies to a detailed, process-oriented investigation of one key area of the proposed framework. In this study the emotional experience of participants (driving force) and its temporal unfolding throughout a learning activity was related to learning. Results showed that a group of students with primarily negative emotional experiences learned the least. Moreover, these students showed an increase in negative emotionality during learning that was predictive of lower learning outcomes. Lastly, additional analyses demonstrated that these emotional processes are related to stable personal dispositions (i.e., trait emotion regulation and neuroticism).

Overall, across all three studies this dissertation has shown that SR shares a common underlying structure across contexts. However, the specific SR processes

required to achieve optimal learning outcomes differ depending on the learning task, context and environment. Through these findings, this dissertation provides a theoretically derived and empirically supported theoretical framework, that situates self-regulated learning within the larger context of self-regulation in education. The findings of the studies are discussed in light of the proposed framework and the added value of a broader conceptualization of SR in education. Key steps for future research programs to extend upon this framework and integrate research traditions on self-regulation in education are derived.

Zusammenfassung

Selbstreguliertes Lernen (SRL) ist eines der wichtigsten theoretischen Konzepte der Bildungsforschung. In Anbetracht aktueller Herausforderungen im Bildungsbereich, wie beispielweise den weit verbreiteten Einsatz von Informationstechnologie in Bildungskontexten, den größer werdenden Fokus Lerner zum lebenslangen Lernen zu befähigen oder dem zunehmenden Schwerpunkt auf Lerner gesteuerte Unterrichtsformate, wird darüber hinaus die zunehmende praktische Relevanz von SRL deutlich. Die Fähigkeit, Lernprozesse effektiv zu regulieren, ist eine Schlüsselfähigkeit für Lernende, um die oben genannten Herausforderungen zu bewältigen. Typischerweise wird SRL als die Regulation und Kontrolle von kognitiven, metakognitiven, motivationalen sowie affektiven Facetten des Lernens zur Erreichung von Lernzielen definiert. Basierend auf dieser breiten Definition haben eine Vielzahl von Forschungsvorhaben SRL aus verschiedenen theoretischen Hintergründen und Perspektiven untersucht. Dabei wurde gezeigt, dass SRL ein zentraler Erfolgsfaktor zur Erreichung von Lernerfolgen in allen Phasen und Bereichen der Bildung eines Individuums ist. Die Vielfalt der Ansätze zur Untersuchung von SRL hat jedoch auch zu Unklarheiten darüber geführt, was SRL ist und wie es am effektivsten gefördert werden kann. Diese Problematik wird noch deutlicher, wenn SRL im Kontext anderer, allgemeinerer Forschungstraditionen zur Selbstregulation (SR) untersucht wird. Die vorliegende Dissertation befasst sich mit dieser Problemstellung. Zu Erreichung dieses Ziels wurden vier Forschungsbereiche zu verschiedenen Aspekten der SR identifiziert und integriert. Diese umfassen Lernaktivitäten (z.B. kognitive und metakognitive Strategien), treibende Kräfte (z.B. Motivation und Affekt), persönliche Dispositionen (z.B. Persönlichkeit) und begrenzte Ressourcen (z.B. Arbeitsgedächtnis und exekutive Funktionen). Auf der Grundlage von starker empirischer Evidenz, die jeden dieser Bereiche eng mit Lernen und akademischen Leistungen verknüpft hat, wurde so ein integratives Rahmenmodell entwickelt, das SRL als Teil von SR in der Bildungskontexten betrachtet. Um dieses Modell empirisch zu testen, wurde in der vorliegenden Dissertation der Vorhersagewert von zentralen, repräsentativen Konstrukten für jeden der Bereiche des Rahmenmodells in verschiedenen Kontexten (z.B. Lernen in der Schule und Lernaufgaben im Labor) getestet. Durch diesen Ansatz ist diese Dissertation die erste

Studie, die die oben genannten Forschungstraditionen zur Selbstregulation im Bildungsbereich empirisch integriert.

Studie I hatte zum Ziel, die besten Prädiktoren für das Lernen in der Schule und für Laborlernaufgaben aus einem umfassenden Satz von selbstregulatorischen Konstrukten zu identifizieren, die die vier im Rahmenmodel postulierten Forschungsbereiche zur Selbstregulation widerspiegeln (Lernaktivitäten, treibende Kräfte, persönliche Dispositionen und begrenzte Ressourcen). Konkret wurden robuste Modelle des maschinellen Lernens verwendet, um die Leistung in der Schule und in Laborlernaufgaben in fünf akademischen Domänen (Mathematik, Physik, Biologie, Kunst und Geschichte) vorherzusagen. Die Ergebnisse zeigten, dass Prädiktoren aus allen Bereichen des vorgeschlagenen Frameworks erforderlich sind, um das Lernen in beiden Settings optimal vorherzusagen. Allerdings unterschieden sich die spezifischen Variablen, die das Lernen in Schul- und Laborlernaufgaben optimal vorhersagten. Während Maße für treibende Kräfte (z.B. Motivation) und begrenzte Ressourcen (z.B. Arbeitsgedächtniskapazität) das Lernen in beiden Settings vorhersagten, zeigten Prädiktoren, die Lernaktivitäten (z.B. Anstrengungs- vs. Wiederholungsstrategien) und Persönlichkeit (z.B. Offenheit) repräsentieren, nur für eines der Lernmaße einen prädiktiven Wert.

Studie II untersuchte, ob und wie sich die Anforderungen an die Selbstregulation bei einer computergestützten Lernaufgabe in Abhängigkeit von der Art der Interaktion der Teilnehmer mit der Lernumgebung unterscheiden. Im Detail nutzten die Teilnehmer entweder mausbasierte oder touchbasierte Interaktion, um mit den Lernmaterialien zu arbeiten. Robuste Modelle des maschinellen Lernens, wurden angewandt, um Lernergebnisse in beiden Bedingungen vorherzusagen. Dazu wurden, ähnlich wie in Studie I, Maße verwenden, die die vier Kernbereiche des vorgeschlagenen Rahmenmodells repräsentieren. Ergebnisse zeigten, dass die Selbstregulationserfordernisse beim Lernen mit Tablets höher waren. Insbesondere wurde das Lernen am Tablet über den Vorhersagewert des Vorwissens hinaus durch kritische Bewertung (Lernaktivität), motivationale Kosten (treibende Kraft), Offenheit (persönliche Disposition) und Task Switching (begrenzte Ressource) am besten vorhergesagt. Leistungsunterschiede bei mausbasierten Interaktionen hingen dagegen nur mit Kontrollmaßen (Leseverständnis und Vorwissen), nicht aber mit selbstregulatorischen Konstrukten zusammen.

Studie III erweiterte das Vorgehen der ersten beiden Studien um eine detaillierte, prozessorientierte Untersuchung eines Schlüsselbereichs des vorgeschlagenen Rahmenmodells. In dieser Studie wurde das emotionale Erleben von Lernen (treibende Kraft) und dessen zeitliche Entfaltung während einer Lernaktivität mit Lernen in Beziehung gesetzt. Ergebnisse zeigten, dass eine Gruppe von Lernenden mit primär negativen emotionalen Erfahrungen am wenigsten lernte. Darüber hinaus zeigten diese Lernenden eine Zunahme negativer Emotionalität während des Lernens, die prädiktiv für geringere Lernerfolge war. Zuletzt zeigten weiterführende Analysen, dass diese emotionalen Prozesse möglicherweise von stabilen persönlichen Dispositionen (Trait Emotionsregulation und Neurotizismus) verursacht werden.

Über alle Studien hinweg hat die vorliegende Dissertation gezeigt, das SR eine zugrundelende Struktur hat, die unabhängig von Kontext ist. Die spezifischen selbstregulatorischen Prozesse, die nötig sind, um optimale Lernergebnisse zu erzielen variieren jedoch nach Rahmenbedingungen (z.B. der Lernaufgabe und -umgebung). Durch diese Studien demonstriert diese Dissertation einen theoretisch abgeleitetes und empirisch gestütztes Rahmenmodell, welches selbstreguliertes Lernen in den größeren Kontext der Selbstregulation in Bildungskontexten setzt. Weitere Schritte für zukünftige Forschungsvorhaben zur Integration von Selbstregulation in Bildungskontexten werden im Kontext des vorgeschlagenen Rahmenmodells hergeleitet und diskutiert.

List of Publications and Contributions

Parts of this thesis were published elsewhere. Following an overview of all publications incorporated in this thesis and my specific contribution to them will be outlined.

Chapter 2: Wortha, F., Brucker, B., Azevedo, R., Tibus, M., Ehlis, AC., Kasneci, E., Nagengast, B., Trautwein, U., & Gerjets, P. (2021). *Robust predictors for learning outcomes in school and laboratory settings: A machine-learning approach for empirically integrating approaches to self-regulation in education*. Unpublished manuscript. University of Tübignen.

The overarching study on self-regulation in education was initially planned, designed, and supervised by P.G., AC.E., E.K., B.N., and U.T. Data collection was organized by M.T. and I assisted in parts of the data collection. The methodology was designed and deployed by me with input from P.G., B.B., and E.K., as well as U.T. and B.N. in early stages of analyses planning. The first draft of the manuscript was written by me. P.G., B.B., and E.K. edited this draft and provided feedback and suggestions. I estimate my contributions to this work to be 80% in total.

Chapter 3: Wortha, F. Mock, P., Brucker, B., Tibus, M., Özbek, O., & Gerjets, P. (2021). *To click or to touch? Learning environments based on tablet versus personal computers differ in the self-regulation requirements imposed onto learners*. Unpublished manuscript. University of Tübignen.

The study was designed by me with input from P.G., B.B., and O.Ö. O.Ö. developed the contents of the learning task with contributions from me, B.B., and P.G. I implemented the experiment. Data collection was organized by M.T. and assisted in parts of the data collection. The methodology was developed by me with contributions from P.M., B.B., and P.G. I further conducted the analyses and wrote the first draft of the manuscript. B.B. and P.G. edited the manuscript and P.M. provided feedback. Overall, I estimate my contributions to be 85% of this work.

Chapter 4: Wortha, F., Azevedo, R., Taub, M., & Narciss, S. (2019). Multiple negative emotions during learning with digital learning environments – Evidence on their detrimental effect on learning from two methodological approaches. *Frontiers in psychology*, *10*, Article 2678

All authors contributed to the conception of the work and revised the final manuscript. RA and MT designed and conducted the study. I conducted the statistical analyses. I and MT wrote the first draft of the manuscript. RA and SN provided several rounds of edits on the manuscript. I estimate my contribution to this work to be 85%.

Table of Contents

Abbreviations

EF	Executive Function
GPA	Grade Point Average
ML	Machine Learning
SR	Self-Regulation
SRL	Self-Regulated Learning
SVM	Support Vector Machine

List of figures

List of tables

List of tables

1. Introduction and Theoretical Framework

"*Give a man a fish he is hungry again in an hour. If you teach him to catch a fish, you do him a good turn.*" (Ritchie, 1885)

According to this well-known English pronoun, teaching individuals how to do something is more beneficial in the long run than doing something for them. At its core this idea mirrors the ideal scenario that researchers, practitioners, and policy makers envision under the term of self-regulated learning (SRL). Rather than teaching students how to solve specific tasks under specific circumstances, students should be enabled to self-sufficiently apply and acquire knowledge, skills, and actions to various (novel) environments and tasks. Therefore, ideally, educational systems should aim to equip learners with a comprehensive set of self-regulatory skills that enable them to adapt to and succeed in any learning situation they are facing throughout their educational and vocational lifespan. While the goal outlined in this idealistic description is very ambitious, it reflects one of the designated objectives in educational systems across the globe. Mayor economic, political, and educational institutions have included self-regulation (SR) and SRL as both major drivers and goals of educational research and practice (National Research Council, 2012; OECD, 2013, 2018; The World Bank Group, 2011). These proposals are based on an extensive amount of research that has shown that student's ability to self-regulate their learning is essential to meet demands, goals, and challenges in contemporary and future education (e.g., Dent & Koenka, 2016; Jansen, Van Leeuwen, Janssen, Jak, & Kester, 2019; Sitzmann & Ely, 2011). Such challenges include, the common use of information technology in educational settings (e.g., Haßler, Major, & Hennessy, 2016), an increased focus on learner-directed learning activities (Hammond & Collins, 2013), and the necessity to maintain life-long learning even after formal education is completed (Field, 2000; Sitzmann & Ely, 2011). The current COVID-19 pandemic has further emphasized most of these issues, as education (temporarily) required distant learning, which comprised the frequent use of technology to engage in learning activities as well as an increased number of learner-directed activities. This imposes additional self-regulatory requirements on learners, that include strategically planning ones learning strategies in (multiple) asynchronous courses, motivating oneself to attend online lectures, conscientiously working on online course materials, and maintaining attention during a video lecture instead of browsing the internet.

An extensive amount of research from varying disciplines has shown that processes related to SRL, and SR in more general, are very effective at mitigating these issues (e.g., McClelland et al., 2018; Schunk & Greene, 2018a). However, the diversity of research united under the umbrella term of SR and SRL, itself poses significant obstacles for researchers and practitioners. Specifically, multiple critical reviews of SR in- and outside of educational research have shown a substantial amount of terminological overlap and confusion between adjacent research fields investigating SR from different perspectives (Inzlicht, Werner, Briskin, & Roberts, 2021; Martin & McLellan, 2008). These articles further have emphasized the need for integrative approaches to connect the broad array of investigations on SR.

This is where the major contribution of this dissertation lies. Specifically, an empirical integration of a multitude of SR constructs in educational settings is introduced. To this end, first, a broad definition as well as core processes and assumptions of SRL shared across theories and models will be identified. Specifically, learning activities representing the cyclical processes of regulation common to all models of SRL are established. Then three adjacent areas of research on SR in education that have been linked to SRL at varying degrees are presented and their connections to learning and SRL investigated. These fields of research were derived from the underlying self-regulatory mechanism they posit and include driving forces (e.g., motivation and affect), personal dispositions (e.g., personality and dispositional interest), and limited cognitive resources (e.g., working memory and executive functions). Based on the identified core processes of SRL (i.e., learning activities) and the adjacent areas of research related to SRL an integrative framework for SR in education is proposed. This framework aims to integrate existing theories of SRL while simultaneously disentangling the processes included in SR during learning. Three empirical studies, that aim to provide empirical evidence for key areas of the proposed framework, will be presented.

1.1 Self-Regulated Learning

Over the last decades SRL has grown into one of the most influential theoretical concepts in educational research (Schunk & Greene, 2018b). Throughout this period, issues related to SRL have been investigated from different angles drawing from diverse theoretical backgrounds, including socio-cognitive (Bandura, 1986, 2001), developmental (Vygotsky, 1978), and metacognitive research traditions (Nelson & Narens, 1994). Based on the various perspectives on SRL a multitude of definitions, theories, and models have been proposed. The conceptual focus of SRL models ranges from metacognitive control processes (Borkowski, Chan, & Muthukrishna, 2000; Winne & Hadwin, 1998), via the interplay of metacognitive and affective processes (Efklides, 2011), to motivational SR processes (Boekaerts, 1996; Pintrich, 2004). Reviews of these approaches have shown that despite the variance in the specific SRL processes covered, core themes, assumptions, and mechanisms are shared between them (Panadero, 2017; Puustinen & Pulkkinen, 2001). However, the extensive diversity of psychological states and processes investigated under the umbrella term of SRL still poses major challenges for future research. While the different approaches to this research topic offer great opportunities for detailed investigation of specific self-regulatory processes (e.g., Greene & Azevedo, 2009, for an example investigating micro-level SRL processes), the sometimes inconsistent terminology and lack of an overarching theoretical framework impedes advances towards a comprehensive understanding of SRL in a larger context (Zeidner, 2019). This issue is further pronounced by the greatly differ levels on granularity of SRL research. Specifically, investigations on SRL range from detailed investigations of micro-level SRL process in controlled laboratory studies through the use of logfile data, to large self-report studies that focus on the use of SRL strategies in school and university. In the following section, the landscape of empirical evidence on the effectiveness of SRL to foster learning and academic achievement will be outlined. Specifically, based on the key findings of meta-analyses on SRL, an overview of the core themes in SRL research will be provided. This synopsis of what has been primarily researched under the umbrella term of SRL will then be used in subsequent sections to derive the functional core of SRL and adjacent fields of research.

1.1.1 Effectiveness of Self-Regulated Learning

SRL is an essential construct in educational research, that is related to key challenges in education (OECD, 2018). Accordingly, an extensive amount of empirical evidence probing the effectiveness of SRL processes in different contexts has been generated. These findings reflect the aforementioned diversity of processes that have been investigated in the context of SRL (e.g., cognitive, affective, metacognitive, and motivational). To summarize this large amount of research and identify focus areas of empirical work as well as factors that moderate the extent of SRLs impact on academic outcomes the focus was exclusively on meta-analyses and meta-analytic reviews that (1) had an international scope, (2) calculated an estimation of the relation between SRL and learning outcomes across multiple studies, and (3) where published in the last 20 years. The criteria were chosen (1) to ensure the findings were representative of general student populations rather than specific (national) educational systems, (2) to focus on reviews that consolidated empirical findings, and (3) to limit the historic scope to findings directly related to the field of SRL research (Zimmerman, 1986) and take into consideration that findings of preceding meta-analyses (i.e., Hattie, Biggs, & Purdie, 1996) have been sufficiently discussed in more recent work (Dignath & Büttner, 2008). This search yielded a total of 11 meta-analyses and meta-analytic reviews (see Table A). These meta-analyses and meta-analytic reviews have summarized the effects of SRL interventions and scaffolds in primary and/or secondary school (de Boer, Donker-Bergstra, Kostons, Korpershoek, & van der Werf, 2012; Dignath, Buettner, & Langfeldt, 2008; Dignath & Büttner, 2008), higher education (R. S. Jansen et al., 2019; Theobald, 2021), work-related trainings (Sitzmann & Ely, 2011), and computer-based learning environments (Zheng, 2016). Similarly, correlational effects of SRL on learning outcomes have been intensely studied in primary and secondary school (Dent & Koenka, 2016), higher education (R. S. Jansen et al., 2019; Richardson, Abraham, & Bond, 2012), simulation-based training (Brydges et al., 2015), as well as web-based environments (Broadbent & Poon, 2015).

Overall, the meta-analyses and meta-analytic reviews found that SRL processes and interventions that aim to foster SRL produced medium to small effects on academic outcomes on average. The beneficial effect of SRL on academic achievement was more pronounced in intervention studies than in correlational investigations. Moreover, SRL has shown beneficial effects for outcomes beyond learning and academic achievement, such as learning strategy use or motivation (Dignath et al., 2008; Dignath & Büttner, 2008; Theobald, 2021).

Several focus areas of SRL research can be inferred on from the predictors, moderators, and mediators included in the meta-analyses (i.e., the content domain, the specific SRL processes investigated, measures of SRL, see Table A). First, the content domain has shown to impact the effectiveness of specific SRL processes. Generally, SRL fostered learning and learning outcomes across a variety of content domains, including math, reading, writing, social sciences, humanities, and medicine (de Boer et al., 2012; Dignath et al., 2008; Dignath & Büttner, 2008; Hoyle & Dent, 2018; R. S. Jansen et al., 2019; Theobald, 2021; Zheng, 2016). However, no consistent pattern of content domains where SRL is most effective was found across these reviews. For instance, in primary and secondary school settings SRL interventions yield stronger effects in math than in reading or other subject domains (Dignath et al., 2008; Dignath & Büttner, 2008), yet correlations between SRL strategy use and academic achievement were stronger in social sciences than in math in similar populations (Dent & Koenka, 2016). For university students, on the other hand, interventions had stronger effects in the field of humanities than in social sciences (Jansen et al., 2019).

Second, with regards to the specific SRL processes the meta-analyses showed that cognitive, metacognitive, and motivational SRL processes are significantly, positively related to learning outcomes. Further, SRL interventions were more effective when they focused on multiple processes and the entire SRL process rather than single SRL processes (Dent & Koenka, 2016; Dignath et al., 2008; Jansen et al., 2019; Zheng, 2016).

Third, it was commonly identified that how SRL and learning outcomes were measured played a moderating role for the effectiveness of SRL. With regard to measuring (learning) outcomes the meta-analyses showed that SRL has an effect on different types of outcomes ranging from posttest in experiments to standardized tests performance and grade point average (GPA) in school and university. However, the effects do not seem to vary in a systematic way, as some studies report the strongest findings for GPA and standardized measures (Dent & Koenka, 2016) while others reported smaller effects for GPA when compared to more proximate measures of performance (e.g., course performance, Jansen et al., 2019; Theobald, 2021). With regard to measuring SRL, results showed that the effects of SRL are particularly pronounced when SRL is captured using online measures (Dent & Koenka, 2016) or count measures for processes (R. S. Jansen et al., 2019) compared to self-reports.

Lastly, some meta-analyses revealed notable factors impacting the effectiveness of SRL, which were not consistently incorporated in all studies. The effectiveness of SRL fluctuated across age and grades level with no consistent systematic pattern (de Boer et al., 2012; Dignath et al., 2008; Theobald, 2021; Zheng, 2016). Moreover, resource management strategies have a significant impact on learning (Broadbent & Poon, 2015; Theobald, 2021), but their impact is small if they are the sole focus (Jansen et al., 2019). Furthermore, the support for the hypothesis that SRL strategies moderate the effect of personal dispositions is limited (Jansen et al., 2019). Finally, SRL interventions seem to be effective regardless of the background of the students (i.e., socioeconomic status, de Boer et al., 2012).

In sum, the empirical evidence for the effectiveness of SRL and interventions that aim to foster SRL showed consistent effects across domains, learning contexts, and outcomes. It was further shown that investigations SRL on SRL typically revolve around the use cognitive and metacognitive learning strategies. Depending on the background of the specific study further processes, such as motivation or resource management strategies are often incorporated. The relative of SRL processes fluctuated unsystematically, indicating that none of the processes was distinctly more important for SR in the researched populations and contexts. Moreover, the meta-analyses have shown that focusing on multiple SRL processes jointly is associated with stronger effects. This indicates that in order to research SRL to its fullest potential, the whole range of SRL and potentially related constructs from neighboring fields need to be considered. However, to pursue this goal, it is first necessary to define the core process of SRL shared throughout empirical and theoretical accounts of SRL. Therefore, in the next section a definitional framework that encapsulates the core processes of SRL will be introduced.

1.1.2 Defining Self-Regulated Learning

As outlined in the previous section the ability to self-regulate one’s learning has a significant positive impact on learning outcomes and academic achievement across tasks and domains. SR in general and SRL in educational context are overarching terms, that encompass multiple psychological constructs and their interactions

(Zeidner, Boekaerts, & Pintrich, 2000). A ground laying definition of SRL referred to learning as self-regulated when learners are *metacognitively*, *motivationally*, and *behaviorally* engaged in a learning activity (Zimmerman, 1986, p. 308). The core of this definition still remains applicable, even though the field of SRL research has made substantial conceptual advances over the years. However, contemporary conceptualizations of SRL have extended upon Zimmerman's definition in several key points. Specifically, SR is described as skills through which learners systematically initiate and maintain cognitive, motivational, behavioral, as well as affective states and processes in pursuit of goals (Schunk & Greene, 2018b). SRL, more specifically, is defined as SR with the goal of learning. Many theoretical perspectives in educational research can be related to SRL as an overarching construct following this definition by Schunk & Greene (2018). These approaches range from investigations of metacognitive monitoring and control during learning (Hacker, Dunlosky, & Graesser, 2009) to research on emotions and regulation in academic settings (Boekaerts & Pekrun, 2016; Harley, Pekrun, Taxer, & Gross, 2019). The diverse SRL models that have been developed based on this diverse background concur with the proposed definition (Schunk & Greene, 2018b), but the weight they assign to individual components of SRL (e.g., cognitive, motivational, behavioral, and affective processes) varies substantially (Puustinen & Pulkkinen, 2001). However, comparisons of SRL models have demonstrated that a set of shared features and mechanisms that unite the different approaches to SRL can be identified (Dent & Koenka, 2016; Panadero, 2017; Puustinen & Pulkkinen, 2001; Schunk & Greene, 2018a). Determining this common core of SR is quintessential to understand how SRL is connected to and distinct from other educational constructs (e.g., motivation and affect, personality, and executive functioning). Thus, the next part will outline crucial principles and mechanisms that characterize SRL based on its most established models.

1.1.3 Core Concepts and Principles of Self-Regulated Learning

Since its inception as a field of research a wealth of well-established and empirically supported theories on SRL have been developed. Detailing their development and mechanisms is beyond the scope of this dissertation and has already been conducted in two theoretical reviews (see Panadero, 2017; Puustinen & Pulkkinen, 2001). Across these reviews seven SRL models have been closely investigated and shared components have been determined. In this dissertation six of these models were used as a basis to derive principal mechanisms and assumptions of SRL. These models include the work of Zimmerman (1989), Boekaerts (1996), Winne & Hadwin (1998), Borkowski (2000), Pintrich (2000), and Efklides (2011). Given that the focus of the present work lies solely in individual learning processes, a recently developed theory focusing on social shared aspects of SR (Hadwin et al., 2011; 2018) was not included in the following examinations. The primary focus of the selected theories includes metacognition (Borkowski et al., 2000; Winne & Hadwin, 1998; Zimmerman, 1989), motivation (Boekaerts, 1996; Pintrich, 2000), interactions of metacognition and affect (Efklides, 2011), and metacognition and motivation (Zimmerman, 1989). However, as outlined above, this categorization may be useful to determine the core component of each model, yet all of these theories incorporate other aspects of SR with differing emphasis. For instance, Borkowski's model of SR (Borkowski et al., 2000) has a clear focus on metacognitive processes. Nonetheless, it also incorporates personal-motivational states (e.g., attributional beliefs). Further, other models have been revised in later work to incorporate additional aspects of SRL that have not been closely considered initially (see Winne & Hadwin, 2008 for an example for motivational processes). This indicates that disentangling SRL into single, isolated processes does not reflect SR during learning appropriately. Instead, central principles and mechanisms applicable to all models have been identified (Puustinen & Pulkkinen, 2001; Schunk & Greene, 2018a).

Figure 1.1 shows a three-phase prototypical process of SRL that can be applied to all selected SRL models. At its core these learning activities span over a preparation, a performance, and an appraisal phase. During a learning episode learners in the first phase, that is the preparation phase, need to analyze the task at hand, set appropriate goals and plan the learning strategies that will be used to achieve these goals (Boekaerts, 1996; Pintrich, 2000; Winne & Hadwin, 1998). Typically, these learning and achievement goals are task specific and set by the learner themself (Boekaerts 1996; Pintrich 2000). The goals set in this preparation phase are essential for subsequent phases of learning as they present a standard to which performance during the learning phase is compared to and which builds the basis for metacognitive monitoring and control (Butler & Winne, 1995; Carver & Scheier, 1998; Nelson & Narens, 1994). Following this preparatory phase, students then enter the performance phase and engage in the planned learning activities (Efklides, 2011; Winne & Hadwin, 1998). During this phase metacognitive monitoring and control are paramount processes. Specifically, learners need to assess if the selected learning strategies are effective and which adaptions might be required to achieve their learning goals (Borkowski et al., 2000; Butler & Winne, 1995; Efklides, 2011; Pintrich, 2000; Winne & Hadwin, 1998; Zimmerman, 2000). These metacognitive judgments are the basis for adjustments of the learning process (i.e., regulation) and can negatively impact learning and achievement when they are inaccurate (Ehrlinger, Johnson, Banner, Dunning, & Kruger, 2008; Koriat, 2012). In the final phase of the SRL processes, learners appraise their learning strategies and performance (Boekaerts, 1996), reflect upon this feedback (Pintrich, 2000; Zimmerman, 2000), and adapt their metacognition (Winne & Hadwin, 1998). Evaluations made in this third phase can have lasting impact on future learning activities. For instance, students might refrain from using a strategy in similar learning tasks in the future if it has yielded unsatisfactory results or set more ambitious learning goals in subsequent tasks if they achieved their learning goals rapidly with little effort. Similar to the performance phase, the adequacy of these adaptions heavily depends on the accuracy of students' appraisals. During SRL activities learners go through multiple cycles of preparation, performance, and appraisal (Winne & Hadwin, 1998, 2008). Some theories further suggest that the phases are rather a loosely coupled sequence of events and that phases might be skipped or students revert to a previous phase without going through the entire cycle (Thillmann, Künsting, Wirth, & Leutner, 2009; Winne, 2001). For instance, students might adjust their goal based on perceived progress during the performance phase before an appraisal of learning outcomes takes place (Butler & Winne, 1995; Carver & Scheier, 1990) or deploy new learning strategies they did not consider in the preparation phase without revisiting their plans. In this context metacognitive monitoring and control become even more important as they build the base for constant adaptions of the learning process throughout all phases of the learning activity (Butler & Winne, 1995, Griffin, Wiley & Salas, 2013, Winne, 2001).

In summary, learning activates describe the prototypical SRL process that includes the core process and assumptions that consolidate SR across different theories and models. This process demonstrates that SRL is (1) an active, innately goal driven process (2) with a cyclical, recursive structure, (3) that is regulated based on metacognitive monitoring and control processes, and (4) includes motivational properties (Schunk & Greene, 2018a). The regulatory processes primarily include cognitive and metacognitive processes at varying levels of granularity (e.g., Greene & Azevedo, 2007, 2009). The effectiveness learning activities further depends on multiple additional factors, such as the learning task and the academic domain (Alexander, Dinsmore, Parkinson, & Winters, 2011; Greene et al., 2015) or the characteristics of the learner (Dörrenbächer & Perels, 2016). Moreover, models and theories of SRL emphasize that, in addition to learning activities, additional processes, such as motivation and affect, are important parts of SR (Boekaerts, 1996; Efklides, 2011; Pintrich, 2000; Schunk & Greene, 2018b; Zimmerman, 2000). Therefore, the following paragraphs will outline three adjacent research areas (driving forces, personal disposition, and limited resources) that are closely related to learning activities and necessary to paint the full picture of SR in education.

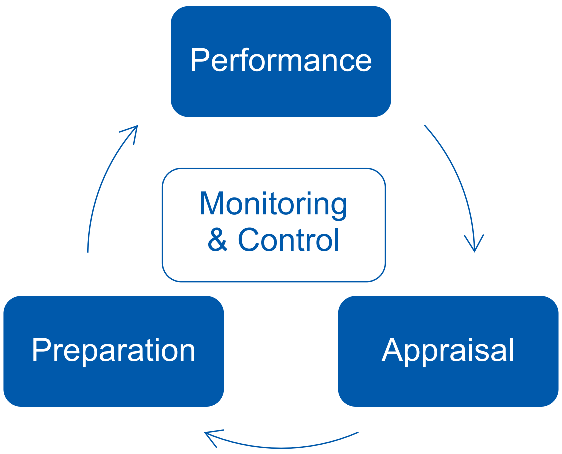


Figure 1.1. The central cyclical process of SRL

1.2 Adjacent Factors of Self-Regulated Learning in Education

The learning activities outlined in the previous section represent the shared core processes across models of SRL. However, SR as a research concept spans far beyond the learning activities outlined in the previous sections. To integrate SRL within a framework of related constructs, it is first necessary to identify and categorize relevant research traditions. There are multiple ways to categorize research on SR in educational contexts, such as the level of granularity, the constructs investigated, or the semantic similarity of the terms used to describe an approach. In this dissertation I decided to focus on the type of mechanism that different research traditions propose. Such differentiations have been previously used to identify the strongest predictors of academic achievement (Richardson et al., 2012). In their review Richardson et al. (2012) have shown that in addition to learning activities the best predictors of learning include, (1) driving forces (e.g., affect and motivation), (2) personal dispositions (e.g., personality), or (3) limited resources (e.g., executive function [EF] and working memory). In the following, sections I briefly outline what these categories represent, their importance in educational research, how they have been incorporated in theories of SRL so far, and which connections to SRL have been established.

1.2.1 Driving Forces

Even though, SRL at its core revolves around cognitive and metacognitive processes (see chapter 1.1.2), the regulation of one's learning is not solely a 'cold' cognitive matter. Initiating and maintaining in learning activities requires forces that energize these learning activities. For example, superior cognitive and metacognitive abilities will not lead to desired learning outcomes if the learner is not motivated to use them (Zimmerman, 2000). Driving forces, however, are not limited to motivational variables, they further include, situation-specific value beliefs and expectancies (Eccles & Wigfield, 2002; Nagengast et al., 2011), situational interests (Hidi & Renninger, 2006), achievement goals (Elliot & McGregor, 2001), or achievement emotions and their regulation (Harley et al., 2019; Pekrun, 2006). Separately, each of these constructs has been linked to learning and academic achievement. For instance, motivation (e.g., expectancy and value, Eccles & Wigfield, 2002) and interventions that aim to foster these motivational components have demonstrated a substantial positive impact on academic outcomes in meta-analyses (Kriegbaum, Becker, & Spinath, 2018; Lazowski & Hulleman, 2016). Moreover, reviews on the effects of emotions on learning have shown that particularly negative emotions, such as anxiety and boredom, have a significant impact on learning (Loderer, Pekrun, & Lester, 2020; Namkung, Peng, & Lin, 2019; Tze, Daniels, & Klassen, 2016). Furthermore, interconnections between many different driving forces in educational settings have been established (e.g., self-efficacy and interest; Huang, 2011, 2016; Rottinghaus et al., 2003).

The majority of theories this dissertation is based on incorporate at least one driving force directly (Boekaerts, 1996; Borkowski et al., 2000; Efklides, 2011; Pintrich, 2004; Zimmerman, 2000). SRL is based on the notion that learners are actively engaged in the learning process (Borkowski et al., 2000; Schunk & Greene, 2018b; Zimmerman, 1986). This active involvement is heavily dependent on learners' interest and motivation for the learning task. For instance, Borkwoski (2000) described a good self-regulated learner (referred to as good information processor) as intrinsically motivated with mastery goals. The importance of driving forces is present throughout the entire SR processes. During the preparation phase, for instance, the goals learners set are dependent on their self-efficacy, interest, and affective associations with the task or domain (Boekaerts, 1996; Efklides, 2011; Pintrich & Zusho, 2002). Performing the learning activities requires that motivation as well as effort are maintained, and that emotions are appropriately regulated (Boekaerts & Pekrun, 2016; Harley et al., 2019; Pintrich, 2003). When learners are faced with impasses during learning it is crucial to overcome the initial negative affective reaction in order to not disengage from the task (D'Mello & Graesser, 2012). Moreover, adaptions of the learning activities, such as increasing goals when a task is not challenging the learner, can serve as means to regulate the learning process (Zimmerman, 2000). In the appraisal phase, driving forces are commonly subject to change. Here, learners update their self-efficacy and other beliefs, goal-orientations, motivation, and attributions based on the outcome of the learning activity (Boekaerts, 1996; Borkowski et al., 2000). Further, affective reactions related to the outcome of the learning process can be elicited. These changes in driving forces will impact subsequent cycles of SRL. For example, underachieving in a learning task might change students goal orientation and lead to avoidance approaches in similar learning activities in the future (Muis & Edwards, 2009). Further, all metacognitive processes throughout the learning activity are subject to affective influences. Through this lens, the impact of affective processes on SRL activities is modelled via their relation to metacognition (cf. Efklides, 2011). For instance, drawing from basic models of SR and affect (Carver & Scheier, 1990, 1998), affect can be elicited through metacognitive monitoring processes. Specifically, as individuals monitor their progress towards goals, they show affective reactions based on the (mis)match of their actual and expected rate of progress. The affective experience in turn can impact attention (e.g., Fredrickson & Branigan, 2005), memory (e.g., Levine & Burgess, 1997), and the accuracy of metacognitive judgments (e.g., Baumeister, Alquist, & Vohs, 2015).

Taken together, driving forces encapsulate a variety of constructs that are closely related to SR, learning, and associated processes. Their role in SRL has been extensively researched since the inception of SRL, primarily in the form of motivational constructs (Zimmerman, 1986) and goal-orientations (Boekaerts & Niemivirta, 2000). More recently, affect and emotions have gained notable attention in SRL studies (Efklides, 2011; Schunk & Greene, 2018a). However, the full breadth of driving forces has not been thoroughly considered in theoretical and empirical work on SRL, despite strong traditions of research in educational settings that have demonstrated the relevance of other driving forces (e.g., interest). Together, both research traditions in- and outside of SRL have clearly shown that processes that energize learning activities are essential to a comprehensive understanding of SRL in educational contexts.

1.2.2 Personal Dispositions

Learning as SRL revolves around monitoring, controlling, and adapting learning processes according to the demands of the current learning task and environment. Although the process-orientation and context dependency are core themes that define SRL, the regulation of learning is also determined by stable characteristics of the learner. In contrast to the other SR processes outlined in this dissertation (i.e., learning activities, driving forces, and cognitive resources) these personal dispositions are primarily defined by their stability over time and situations. Specifically, they involve patterns of thoughts, feelings, and behaviors that, in light of their stability over time, provide more distal causes of SRL processes. Prominent dispositions in the context of SRL include personality traits (e.g., conscientiousness, Costa & McCrae, 2008) as well as other stable social cognitive constructs such as academic self-concepts and dispositional interests (Hidi & Renninger, 2006; Marsh, 1990), mindsets (David Scott Yeager & Dweck, 2012), or cognitive abilities (e.g., Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007). Even though, these constructs are characterized by temporal stability it is important to note that they may develop and change over the life-span, particularly during childhood and adolescence (Roberts, Walton, & Viechtbauer, 2006). There is further evidence that the development of personality traits continues throughout later stages of life (Roberts & Mroczek, 2008). In educational contexts research on personal dispositions has primarily focused on conscientiousness (one of the big five traits, Costa & McCrae, 1998), on grit (Duckworth, Peterson, Matthews, & Kelly, 2007), or on a combination of personality and stable motivation (Trautwein et al., 2015). Empirical evidence underlined the importance of these constructs in education as they are robust predictors of academic success (mediated by academic effort) that are fairly independent of cognitive ability (De Raad & Schouwenburg, 1996; Poropat, 2009; Rimfeld, Kovas, Dale, & Plomin, 2016).

In theories and models of SRL the role of personal dispositions is less clearly outlined. Generally, it is postulated that SRL processes moderate the effect of personal dispositions on learning outcomes (Schunk & Greene, 2018b). Yet, the incorporation of dispositions varies substantially between models. Some models have included unspecified dispositions indirectly in the concept of ‘self’ (Boekaerts, 1996; Zimmerman, 2000). In Winne and Hadwin’s (1998) model of SRL, personal beliefs, dispositions and styles are part of cognitive conditions that the learner brings to the learning task. The primary way of personal dispositions to impact the learning process is realized through their direct relation to the standards used as basis for monitoring control. The metacognitive and affective model of SRL (Efklides, 2011) describes SR on two levels – the person level and the task x person level. Here dispositions are situated at the person level and include attitudes and affective dispositions. Moreover, this model contains explicit assumptions regarding the changes in importance between the levels throughout a learning activity. Specifically, it suggests that with ongoing learning processes the focus heavily shifts toward the task x person level and bottom-up aspects of SR. Taken together with the finding that individuals who show high levels of trait SR often not engage in in-situ SR as commonly as they can effectively avoid unfavorable situations (Hill, Nickel, & Roberts, 2014; Hofmann, Schmeichel, & Baddeley, 2012; Inzlicht et al., 2021), this indicates that personal dispositions might not demonstrate close relations to task-specific measures of learning in some cases.

With regard to empirical connections between SRL and dispositions, only few studies have provided empirical evidence. They have shown that conscientiousness was positively related to motivational aspects of learning, self-reported SRL strategy use, and achievement (Chamorro-Premuzic & Furnham, 2003; Eilam, Zeidner, & Aharon, 2009). Moreover, conscientiousness and openness were found to be associated with a more frequent use of metacognitive and elaborative learning strategies (Bidjerano & Dai, 2007). Neuroticism on the other hand had a negative relation to achievement as it is associated with emotional instability (Chamorro-Premuzic & Furnham, 2003).

In sum, personal dispositions provide distal explanations that can shed light on how students do self-regulate across learning tasks and longer time frames (e.g., a whole semester etc.). With regard to SRL activities they provide a basis to explain if and how students generally approach and engage in tasks or show predispositions for certain behaviors, cognitions, and affects (e.g., neuroticism as trait vulnerability to negative emotions). Further, in order to bridge the gap between different research traditions on SR across levels of granularity (e.g., laboratory learning tasks vs. academic achievement in schools), it is essential to include personal dispositions into an integrative framework of SR in education.

1.2.3 Limited Resources

Actively engaging in SRL processes is a challenging and effortful endeavor, which, as a consequence of its cognitively demanding nature, requires considerable cognitive resources (Pintrich & Zusho, 2002; Winne, 1995). High-level learning-content-related (e.g., writing a summary) and learning-process-related (meta-)cognitive operations (e.g., monitoring the learning process) encompassed in SRL are based on different kinds of limited processing resources, particularly working memory and executive functioning. Similar to SRL, research on EF has been approached from multiple angles, resulting in numerous definitions and concepts of EF (Suchy, 2009). These range from conceptualizations of EF that are focused on updating, inhibition, and shifting (Miyake et al., 2000), which have been more recently integrated into a hierarchical structure where inhibition is subsumed by a common EF factor (Miyake & Friedman, 2012), to distinctions between metacognitive and emotional (Ardila, 2008) or ‘hot’ and ‘cool’ EFs (Zelazo & Carlson, 2012). A commonly agreed upon definition describes EF as domain-general, higher-level cognitive processes, that, through their influence on lower-level cognitive functions (e.g., visual attention), enable individuals to regulate thoughts and actions in service of goal directed behaviors (Friedman & Miyake, 2017; Miyake et al., 2000). In other words, EF are cognitive resources that are required to represent and choose between goals and strategies (Miller & Cohen, 2001), which is quintessential for goal directed behavior, such as SRL. Moreover, EF and working memory provide a basis for cognitive control and metacognition (Alloway & Alloway, 2010; Cowan, 2014; Miyake et al., 2000), which, in turn, are at the center of any SRL activity. Beyond EF and working memory, further attentional and volitional resources can be potentially required to ensure goal achievement (e.g., for effort investment, regulation of emotions, and the willpower to resist impulses; Bauer & Baumeister, 2011; Gross, 2013; Mischel et al., 2011).

Yet, theories of SRL on the other hand haven’t addressed these underlying cognitive resources comprehensively. Only Borkowski’s (2000) process model of metacognition explicitly mentions EF as an integral component of SRL. In this model, based on Butterfield’s theory of EF (Butterfield, Albertson, & Johnston, 1995), monitoring and control are defined as EFs. These EFs further consist of three main components: task analysis, strategy control, and strategy monitoring. According to Borkowski’s model (2000), EFs build the bridge between motivational states and (learning) strategy knowledge, contextualized in the current learning task. Other models, on the other hand, solely imply that cognitive resources are necessary for effective regulation (Pintrich, 2000; Winne & Hadwin, 1998) without specifying them in detail. This sparse incorporation of underlying cognitive resources in theories and models of SRL is in stark contrast with a plethora of empirical evidence demonstrating the importance of cognitive resources for learning and academic achievement in educational research (Alloway, 2006; Alloway & Alloway, 2010; Cirino & Willcutt, 2017; Diamond, 2013; Hutchins, Wickens, Carolan, & Cumming, 2013; Paas & Sweller, 2012; Xie et al., 2017). Firm relations between working memory, executive functions, and academic achievement have been shown in different educational contexts and populations (Alloway & Alloway, 2010; Cirino & Willcutt, 2017; Diamond, 2013). In research on learning and instruction, the impact of cognitive resources, particularly working memory, is most commonly researched through the lens of cognitive load theory (Paas & Sweller, 2012). In this theoretical framework the cognitive load (the amount of working memory used) that certain learning task and environment characteristics evoke is investigated under the assumption that learning is most effective when learners are experiencing an optimal level of cognitive load (Castro-Alonso, de Koning, Fiorella, & Paas, 2021). Meta-analyses have shown that designing learning tasks and interventions to reduce cognitive load is an effective way to foster learning (e.g., Hutchins et al., 2013; Xie et al., 2017).

Based on the importance of cognitive resources in educational contexts outside of SRL research, scholars have recently begun to integrate SRL and cognitive resources (e.g., drawing links between cognitive load theory and SRL; de Bruin & van Merriënboer, 2017). With regards to EF, first studies have shown that EF and SRL are closely related, particularly with regard to metacognitive processes (Effeney, Carroll, & Bahr, 2013; Follmer & Sperling, 2016; Rutherford, Buschkuehl, Jaeggi, & Farkas, 2018). Particularly, measures of shifting and inhibition were linked to SRL and academic achievement (Follmer & Sperling, 2016; Rutherford et al., 2018). Taken together these studies suggest that EF and other cognitive resources can support SRL but also form a congestion in the SRL processes that may limit the unfolding of other aspects of SR (e.g., working memory constraining SRL processes, Pintrich & Zusho, 2002).

All in all, the theoretical and empirical contributions outlined above show that cognitive resources need to be considered in an integrative framework of SRL to accurately depict SRL as goal driven processes. Despite the surprising scarcity of considerations of cognitive resources in SRL research so far, research inside and outside of educational settings has clearly shown the importance of EF, working memory, and related constructs are essential to understand the unfolding of SR processes.

1.3 Integrating SRL into a Framework for Self-Regulation in Education

In the previous sections I have outlined that SRL in the form it has been researched till now is not an isolated and self-sufficient field of research. Instead, research on SRL draws from and is potentially influenced by many neighboring fields of research in educational settings. However, the previous chapters have also shown, that the different research traditions related to SR in educational settings mostly co-exist rather than being systematically interrelated. This disintegrated state of research has led to conceptual and terminological issues. Most commonly these problems are categorized as jingle and jangle fallacies. Jingle fallacies are defined by different things bearing the same name. A prominent example in the context of SR is the frequently interchangeable use of SR and self-control. The importance of differentiating SR from self-control and other similar constructs has been pinpointed repeatedly (e.g., Inzlicht et al., 2021; McClelland et al., 2018). A jangle fallacy on the other hand describes the same construct being referred to under different names. A prominent example with high relevance to educational research are studies on grit (Duckworth et al., 2007) and conscientiousness (Roberts, Jackson, Fayard, Edmonds, & Meints, 2009). Recently the conceptual overlap between these constructs have been extensively investigated and have shown that grit and conscientiousness are not separable constructs (Muenks, Wigfield, Yang, & O'Neal, 2017; Ponnock et al., 2020). Both fallacies negatively affect the integration of research on SR in- and outside of educational settings. Specifically, using different names for the same construct (e.g., grit and conscientiousness) leads to an overestimation of the amount of relevant self-regulatory constructs for learning. In this case literature on a surface level suggests that both grit and conscientiousness are separate predictors of learning, when they describe the same influence factor for learning processes. This could in turn lead to an underestimation of the importance of the underlying construct. For instance, conscientiousness' importance for learning could be underestimated in cases when literature on grit (which uses a different name for same construct) is not considered, as findings for grit are not attributed to the underlying trait conscientiousness. On the other hand, a jingle fallacy (e.g., the use of SR when self-control is researched) creates the illusion that research on vastly different constructs is directly related. Thus, the erroneous assumption could arise that SR(L) is relevant for outcomes that have only been researched in the context of self-control. But through the interchangeable use of both constructs these findings are also attributed to SR. Therefore, both fallacies present considerable obstacles for integrative research approaches on SR in education. This dissertation provides insights to remedy such issues by classifying constructs related to SR based on the underlying mechanism they proposed. By differentiating constructs based on how they are supposed to effect SR and learning, similar or identical constructs are categorized. As outlined in previous sections, in the proposed framework the focus will be on four types of constructs: learning activities, driving forces, personal dispositions, and limited resources.

1.3.1 An Integrative Framework of Self-Regulation in Education

In the previous sections I have outlined how different research traditions investigating driving forces, personal dispositions, and limited resources are related to learning and SRL specifically. Based on these relationships a framework was developed that situates SRL in a larger context of self-regulatory constructs in educational settings. As illustrated in Figure 1.2, according to this framework SR in learning situations unfolds as the interplay of learning activities, driving forces, and limited resources. Specifically, learning activities describe the cyclical processes that is shared across theories of SRL (preparation, performance, and appraisal see chapter 1.1.2). The focus in these activities lies in the use of cognitive learning strategies and metacognitive processes that are used to control and adapt these strategies. To effectively engage in and maintain these activities driving forces (e.g., motivation, affect and their regulation, see chapter 1.2.1) and (limited) cognitive resources are required (e.g., working memory and executive functioning, see chapter 1.2.3). Similar to models of SRL that focus on motivational and affective aspects (Boekaerts, 1996; Efklides, 2011; Pintrich, 2000; Zimmerman, 2000), potential influences of such 'hot' non-cognitive processes are considered throughout all phases of the learning activity. Moreover, through the inclusion of cognitive resources the proposed framework closes a gap in SRL research. These underlying low-level cognitive processes and resources build the basis for the high-level learning activities. Furthermore, cognitive resources can interact with driving forces, for instance in cases where negative affective reactions bind or alter cognitive resources (e.g., Kensinger & Corkin, 2003). The interplay of learning activities, driving forces, and cognitive resources as SR process is situated within the learning task, environment, and context. In other words, the unfolding of the three core components of SR is context specific. For instance, studies have demonstrated that the effectiveness of learning strategies (Alexander et al., 2011; Greene et al., 2015) as well as the intensity of affective reactions (Boekaerts & Pekrun, 2016; Namkung et al., 2019) vary across content domains. Research on cognitive load in educational contexts has further shown that features of the learning task and domain directly affect the available working memory capacity (Paas & Sweller, 2012; Sweller, 2005; Van Merrienboer & Sweller, 2005). Personal dispositions on the other hand, influence the SRL processes across contexts. In this model they represent stable patterns of thoughts and behaviors, that determine which and how processes typically unfold. The impact of these processes on learning outcomes is always moderated through the SR processes in the specific learning situation. For instance, trait vulnerability to negative emotional experience (i.e., neuroticism) may prime driving forces towards pronounced negative reactions when impasses are encountered. However, if no such hurdle is present or scaffolds are provided, which assist the learner to overcome these issues (D'Mello & Graesser, 2013), this personal trait will not affect the learning outcome in that situation. In the proposed framework these outcomes include learning and academic achievement, as well as affective and motivational outcomes. Outcomes of the learning process in turn can lead to changes in dispositions or the self-regulatory processes, particularly if a larger learning task requires multiple SR cycles. Lastly, the SR cycle in a given learning situation can feedback into personal dispositions and lead to changes, for instance, strenuous SR processes might have long-term negative effects on dispositional motivation and self-efficacy.

The main purpose of this framework is to situate existing research, models, and theories of SRL in a broader context rather than proposing new mechanisms. Existing models of SRL can be situated within this framework and their processes can be interpreted in a larger context of SR. For instance, the detailed metacognitive processes postulated by Winne & Hadwin (1998) and their extension in Greene & Azevedo (2009) are situated close to the learning activities themselves. The cognitive conditions they propose are cognitive resources (e.g., prior knowledge) and personal dispositions (e.g., beliefs and dispositions). Boekaerts (1996) two pathways of SR, for instance, take place at the intersection of driving forces and learning activities. The person and person x task levels introduced by the metacognitive and affective model of SRL (Efklides, 2011) span from personal dispositions, via driving forces to learning activities.

Beyond the incorporation of existing models of SRL, potential interactions between different areas of SR are of particular interest. The processes involved in this broader framework are interdependent to the extent that a minimum activity in each of the areas is required for a SRL activity to unfold. A central assumption of SRL processes outlines that learners need to be actively engaged in the learning process. Hence, sufficient extrinsic or intrinsic motivation to pursue the learning task is required (Borkowski et al., 2000; Schunk & Greene, 2018b). However, a motivated student might not achieve good learning outcomes when they do not engage in adequate learning activities (e.g., engage in inappropriate learning strategies; failure to adjust learning strategies due to poor metacognitive monitoring accuracy). Furthermore, even if the most appropriate strategies are known and engaged in, success is not guaranteed if the student does not have the required cognitive resources to properly execute these strategies and deeply process the learning materials (e.g., experience of overwhelming cognitive load; reduced working memory capacity).

As indicated above the importance of personal dispositions can be determined by other aspects of the framework. For instance, conscientiousness as a personal disposition and interest, a situational driving force, may show compensatory patterns. The trait conscientiousness is only required to bring about sufficient invested effort if learners lack interest in the current content domain and learning task (Trautwein et al., 2015). However, if sufficient interest is present, academic effort will be invested regardless of personal dispositions. This effort in turn is likely to lead to the use of learning strategies and cognitive resources to achieve high learning outcomes.

Another interaction potentially takes place at the intersection of driving forces and limited resources. While the potential effect of emotions on metacognitive judgements and accuracy has been previously shown (e.g., Baumeister et al., 2015), the effect of driving forces on cognitive resources is rarely considered despite research that indicated a potential negative effect on working memory (e.g., Kensinger & Corkin, 2003). This indicates that negative driving forces (e.g., negative emotions) can impact the SR process in multiple ways.

Overall, the proposed framework brings together the fields of research on SR identified in this dissertation. Through considering existing theories and models of SRL, including their core mechanism and simultaneously disentangling further SR processes to which they are related to, existing research can be situated in a larger framework. Moreover, key families of variables that need to be considered in integrative studies of SR and their potential interactions are detailed. The empirical studies in this dissertation will showcase empirical evidence for key components of this framework in order to demonstrate the added value of investigating SRL in the larger context of SR in education.

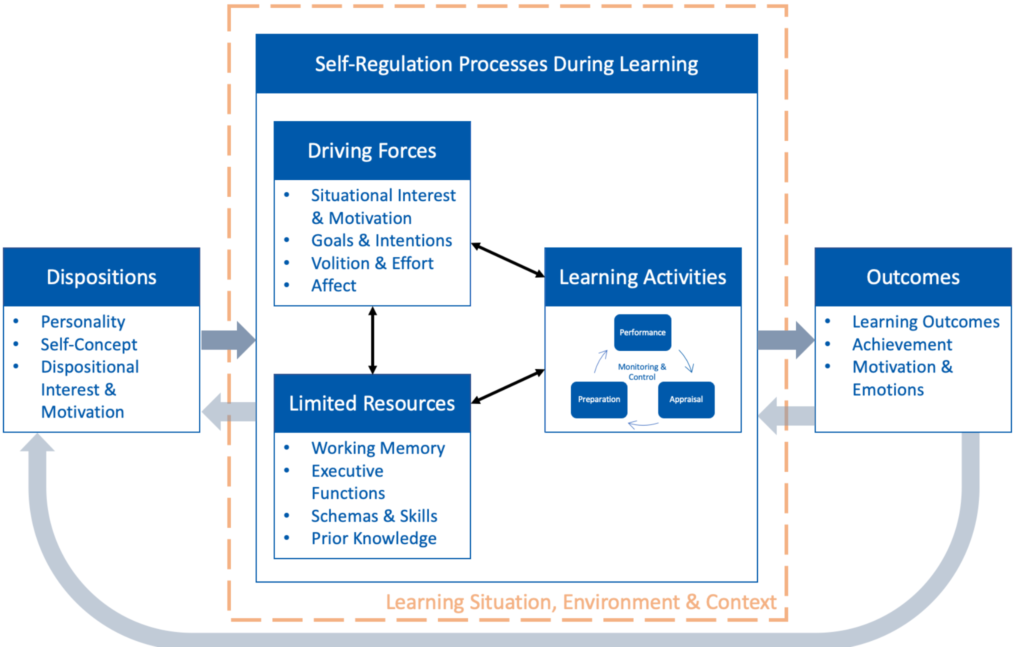


Figure 1.2. *An integrative framework for self-regulation in education.*

1.3.2 Challenges of integrating Self-Regulated Learning into a larger framework

Integrating the diverse research traditions outlined above poses major challenges. In addition to the definitional and conceptual issues, major methodological hurdles have to be overcome to provide an empirical integration of SRL into the larger context of SR. To adequately represent the different research traditions, a variety of measures ranging from (neuro-)cognitive laboratory tasks that often measure reaction times at a millisecond scale (e.g., Stroop test, Miyake et al., 2000), to questionnaires that assess students' stable dispositions (e.g., personality, Costa & McCrae, 2008) need to be integrated. Moreover, in the context of learning, these measures are in turn related to different outcome measures, including performance in experimental learning tasks (e.g., Azevedo & Cromley, 2004) or standardized measures of achievement

(Duckworth, Tsukayama, & May, 2010). Collecting such extensive amounts of data at different levels of granularity for sufficiently sized samples is typically not feasible. Furthermore, empirically integrating constructs across fields of research introduces major methodological challenges. Relations between variables from different fields of research are often unstable, due to differences in measurement properties (Dang, King, & Inzlicht, 2020). For instance, a recent study that integrated different SR constructs in a clinical setting through data driven methods found that survey data and cognitive tasks showed little relations, but within each of these measures structures of SR constructs were identifiable (Eisenberg et al., 2019). This shows that an empirical integration of SRL and related constructs in education requires substantial analytical effort to reliably reveal interconnections between different research traditions. A particularly promising methodology in this context is the use of machine learning. Machine learning approaches show great potential to address methodological issues and extend the scope of psychological research (Yarkoni & Westfall, 2017), including clinical (Dwyer, Falkai, & Koutsouleris, 2018) and educational questions (Hilbert et al., 2021). A specific advantage of machine learning approaches is their ability to effectively address issues of traditional statistical analyses, particularly when handling large numbers of variables. For instance, the widely used correction of p-values in multiple testing scenarios has been shown to not sufficiently reduce the false positive rate when dealing with many predictor variables (Eklund, Nichols, & Knutsson, 2016). Further, overfitting is a frequent problem in psychological research and related disciplines, resulting in models that primarily reflect noise and special properties of a specific data set (Dwyer et al., 2018; Whelan & Garavan, 2014). This may yield results that cannot be reproduced, which, in the light of a lack of replicability of psychological findings has raised major concerns regarding scientific and analytical practices (Ioannidis, 2005; Maxwell, Lau, & Howard, 2015). In the field of machine learning a variety of countermeasures have been developed to address these particular issues. These include cross-validation approaches to avoid overfitting, a stronger focus on (non-linear) patterns rather than single predictors, robust feature selection and dimensionality reduction methods, and generalization (i.e., the performance on unseen data) as a key outcome (Dwyer et al., 2018). The utility of machine learning approaches in educational contexts goes far beyond the provision of novel analytical techniques (Hilbert et al., 2021), as they can broaden the scope of research to address questions that traditional analytical approaches cannot address. For example, machine learning predictions can be used to identify struggling students in early stages of a course (Đambić, Krajcar, & Bele, 2016), which can serve as basis for timely interventions.

For this dissertation machine learning approaches are particularly promising as each of the questions addressed incorporates a substantial amount of potential self-regulatory constructs that need to be analyzed simultaneously in order to empirically integrate different traditions of SR research in education (see Study I and Study II). Therefore, robust modelling approaches to select appropriate predictors or identify an underlying structure in predictor variables were deployed at different levels of granularity. In addition, to robust generalizable predictions, the focus further lied in explainable models. Revealing and interpreting the relation(s) between different self-regulatory constructs and learning is a crucial step in achieving the overall goal of this dissertation to situate SRL with a larger framework of SR in education.

1.4 Research Questions

1.4.1 Objective of this Dissertation

The aim of the current thesis is to empirically extend the scope of research on SRL and take a first step towards an integrative theory of SR in educational settings. More specifically, the core focus of SRL research, which is centered around the regulation of learning activities and behaviors through metacognitive monitoring and control (chapter 1.1.2), is expanded upon by driving forces that energize the learning process (chapter 1.2.1), personal dispositions that represent relatively stable tendencies and habits (chapter 1.2.2), and limited resources which are required to carry out high-level cognitive activities (chapter 1.2.3). Although these research traditions showed close relations to learning and can be directly related to self-regulatory processes in educational settings, attempts to integrate them are still missing in research. Besides of investigations that aimed to build the bridge between two adjacent fields (e.g., Bidjerano & Dai, 2007; de Bruin & van Merriënboer, 2017; Efklides, 2011), research on SR in education remains in fragmented state. To empirically address the disintegrated nature of co-existing research traditions, the present dissertation analyzed a comprehensive selection of self-regulatory measures, ranging from performance data in cognitive tasks to questionnaires measuring personality. More importantly, I particularly focused on the relation of self-regulatory constructs to learning outcomes across different levels of granularity.

Outside of educational research, first attempts to integrate multiple approaches to SR have been conducted on a conceptual (Inzlicht et al., 2021) and empirical level (Eisenberg et al., 2019). These investigations have shown that unifying self-regulatory constructs is a necessary step towards a comprehensive understanding, but is conceptually and methodologically challenging (Dang et al., 2020). Therefore, this dissertation focused on robust modeling approaches to counteract potential issues related to complexity and granularity of data from diverse theoretical backgrounds. Through machine learning predictions and a combination of person- and variable-centered statistical approaches the studies in this dissertation showed how the complexity of data can be reduced across levels of granularity and important patterns can be identified in order to obtain reliable, generalizable, and interpretable results.

1.4.2 Overview of the Studies in this Dissertation

Three studies were conducted to answer three key questions aiming at integrating approaches to SRL (see Figure 1.3).

The first study addressed a central educational question with high scientific and practical value. What are the best predictors of learning? Further it investigated if these characteristics are the same for schools and laboratory learning tasks. With regards to the fragmented state of research, comparing the predictive value of SR processes in both settings can provide indications which aspects of SR are generally important for learning and which are more specific to certain contexts. To this end, data collected in a unique and extensive study on SR in educational contexts was used. In this study, over 300 participants filled in questionnaires, completed cognitive tasks, and learned about five different topics in laboratory learning tasks. I included predictor variables that represent each of the four core areas of the proposed framework evenly and used them to predict the average grade at the end of high school across five domains and the average performance in five corresponding learning tasks. Due to the large number of predictors used and the exploratory nature of this question, machine learning models were developed and used. A robust feature selection procedure coupled with parsimonious prediction algorithms ensured that results were generalizable and interpretable. The findings in this study were a key step towards revealing and understanding the added value of multiple aspects of SR in educational research.

The second study narrows the integrative approach of the first study down to investigations to a specific contextual variable - the learning medium. Theories of SRL emphasize that learning is a highly context-dependent process (Greene et al., 2015; Pintrich, 2000; Winne & Hadwin, 2008; Zimmerman, 2000). Accordingly, the predictive value of self-regulatory variables is likely to vary across contexts. One such environmental factor, that has gained immense intention through recent developments in education, are touch-interactions. Specifically, touch-devices (e.g., tablet PCs) have become a commonly used technology in schools and universities. Research has shown that tablets have great potential to foster different aspects of SR, but may also impose challenges on learners (Mulet, Van De Leemput, & Amadieu, 2019). Specifically, learning with tablets can affect driving forces (Mulet et al., 2019), learning activities (Sidi, Shpigelman, Zalmanov, & Ackerman, 2017), and cognitive resources

(Abrams, Weidler, & Suh, 2015). In this study, I used data for a multi-perspective hypermedia art-learning task from the same data set as Study I. The measures investigated were comparable to Study I, but task and domain specific variables were used where applicable (e.g., interest in art, prior knowledge and metacognitive accuracy in the learning task). Potential differences in the predictive value between participants who learned with a PC or tablet were analyzed through robust machine learning classifications. Findings from this study demonstrated how contextual factors effect multiple aspects of SR.

The final study of this dissertation, further focused on a more-detailed level of granularity. While the first two studies focused on SRL processes in a state-like form (i.e., though self-reports at a singular time-points), this study focused on the unfolding of one particular aspect of SRL during a learning activity and its relation to students' dispositions. Specifically, I investigated the unfolding of affective processes during learning with MetaTutor, a hypermedia learning environment covering biological contents (Azevedo, Johnson, Chauncey, & Burkett, 2010). This study addressed the issues, that driving forces, particularly emotions, are conceptualized as dynamic processes in the context of SRL (e.g., D'Mello & Graesser, 2012; Harley et al., 2019), but their temporal dynamics and interactions are often not sufficiently addressed. Further, similar to SR, emotions during learning have been researched from many angles (D'Mello & Graesser, 2012; Ekman, 1992; Pekrun, 2006), yet, connections between different approaches to emotions are still rare. A triangulation of person- and variable-centered approaches was used to identify learners with comparable emotional experiences, reveal a stable low-dimensional structure of emotions, and investigate their development over time and interaction with learning. Moreover, to investigate the effect of personal dispositions in this context, the study investigated if emotional dynamics during learning and their effect on learning were related to habitual emotion regulation strategies. To further test the proposed mediation of personal disposition via SRL processes, additional analyses addressed if trait neuroticism (Costa & McCrae, 2008) is related to differences in emotional experience during learning.

Together these three studies aimed to showcase the utility of the proposed framework for research on SRL. Across levels of granularity, from investigations of learning activities, driving forces, personal dispositions, and cognitive resources across learning tasks and school subjects to detailed analyses of the unfolding of specific driving forces in a laboratory tasks and their relation to personal dispositions, the key components of the framework were substantiated with empirical evidence. The specific aspects of the framework addressed by each study are displayed in Figure 1.3.

	Question	Areas of SRL	Learning Task(s)	Methods
1. What are the best predictors of learning in school and in laboratory learning tasks?				
Study I	Is learning in school and in the lab defined by the same or different SR predictors?	Learning Activities Dispositions Driving Forces Cognitive Resources	Grades and laboratory tasks in math, physics, biology, arts, and history	Robust machine learning predictions
2. Does the predictive value of self-regulatory variables differ by study medium?				
Study II	Does the study medium affect the predictive values of SR predictors?	Learning Activities Dispositions Driving Forces Cognitive Resources	Multi-perspective-hypermedia learning in arts	Robust machine learning predictions
3. How do emotional self-regulated learning processes predict learning?				
Study III	How are specific combinations of emotions related to SRL?	Driving Forces Dispositions[1]	Hypermedia learning in biology	Person- and variable centered clustering Inferential statistics

Figure 1.3. *Overview of the studies in this dissertation. SR: self-regulation, SRL: self-regulated learning*

Note. [1]Effects of personal dispositions were operationally defined as habitual emotion regulation strategies in this study. Dispositions in line with the previous studies (i.e., personality) were investigated in additional analyses for this dissertation (see Appendix B).

2 Study I

Robust predictors for learning outcomes in school and laboratory settings: A machine-learning approach for empirically integrating approaches to self-regulation in education.

Franz Wortha, Birgit Brucker, Roger Azevedo, Maike Tibus, Ann-Christine Ehlis, Enkelejda Kasneci, Benjamin Nagengast, Ulrich Trautwein, Peter Gerjets

Note. Unpublished Manuscript. This article might not exactly replicate the final version published in the journal.

Estimated contributions

Scientific ideas by the candidate (%)	Data generation by the candidate (%)	Analysis and Interpretation by the candidate (%)	Paper writing done by the candidate (%)
60	10	90	90

Abstract

Self-regulation is a core success factor in many aspects of life (e.g., health care, economics, and education). Especially in the field of education the ability to self-regulate has become a focal point of research and practice as it presents an essential skill that determines learning outcomes and academic achievement across different contexts (e.g., in school and laboratory learning tasks). The strong interest in self-regulation in many disciplines of educational research has led to an influx of different constructs investigated under the umbrella term of self-regulation. This resulted in a fragmented state of research, which imposes significant challenges for researchers and practitioners. This study aimed to address this issue by empirically integrating central areas of research on self-regulation in education. Specifically, based on their underlying mechanism we identified and investigated four groups of self-regulatory constructs: learning activities (e.g., cognitive and metacognitive strategies), driving forces (e.g., motivation and affect), limited resources (e.g., working memory and executive functions), and personal dispositions (e.g., personality) and their relation to learning. To this end, a comprehensive set of measures representing these four research traditions was collected for 321 university students. Robust machine learning models were used to predict school grades and performance in laboratory learning tasks across five academic domains. Results showed that optimal predictions of both outcomes required predictors from all research traditions. However, the specific predictors that best predicted performance varied between outcomes. This indicated that self-regulation at different levels of granularity (i.e., school grades and performance in laboratory learning tasks) shares the same underlying structure (i.e., learning activities, driving forces, limited resources, and personal dispositions) but the specific self-regulatory requirements are situation specific. Implications for further steps to integrate self-regulation in education were discussed.

Introduction

Self-regulation entails a quintessential set of skills for success across all areas of human behavior. The ability to effectively self-regulate has been repeatably linked to important life outcomes, including health, well-being, and educational achievement (Baumeister, Heatherton, & Tice, 1994; Burnette, O'Boyle, VanEpps, Pollack, & Finkel, 2013; Dignath, Buettner, & Langfeldt, 2008; Moffitt et al., 2011; Robson, Allen, & Howard, 2020; van Genugten, Dusseldorp, Massey, & van Empelen, 2017). Yet, the scope, theories, and models used under the umbrella term of research on self-regulation greatly differ. Scholarly investigation in this context span from neuro-scientific investigations of the neural correlates of self-regulatory processes on a millisecond level to research in personality psychology on long-term goal-oriented behaviors that extent over years. The terminology related to self-regulation within and between disciplines shows significant inconsistencies, which has been criticized for years (e.g., Duckworth & Kern, 2011). But also, very recently researchers from different disciplines, including cognitive, personality, and clinical psychology, have showcased again that the current state of research still remains fragmented and that this poses major challenges in advancing and integrating the theoretical understanding of self-regulation (Eisenberg et al., 2019; Inzlicht, Werner, Briskin, & Roberts, 2021; Malanchini, Engelhardt, Grotzinger, Harden, & Tucker-Drob, 2019; Nigg, 2017).

In education self-regulation is particularly important. It is seen as a central set of skills for students to meet the requirements and challenges in modern education, including the increased use of advanced information technology, the growing focus on learner-controlled activities, and the expanding emphasis on lifelong learning (National Research Council, 2012; OECD, 2013; World Bank 2011). In the context of these educational challenges self-regulation refers to processes that involve planning, monitoring, controlling and reflecting upon thoughts, behaviors, motivation, and emotions in pursuit of goals (Schunk & Greene, 2018). In consonance with this broad definition a variety of regulatory processes have been directly linked to learning and learning outcomes. For instance, research investigating self-regulated learning has shown that effective deployment, monitoring, and control of cognitive, affective, metacognitive, and motivational processes are key success factors for learning in school and more controlled laboratory settings (e.g., Azevedo et al., 2018; Dignath &

Büttner, 2008). Furthermore, researchers focusing on personality and other stable dispositions demonstrated that conscientiousness and similar constructs (e.g., grit) are powerful predictors for academic outcomes (e.g., Muenks et al., 2017; Roberts et al., 2007). Scholars that research different aspects of executive functioning and working memory, in turn, found that these cognitive abilities and resources are good predictors of academic outcomes even when controlling for general mental ability (e.g., Alloway & Alloway, 2010). While some studies have connected constructs from multiple research traditions, such as for example personal dispositions, and learning strategy use to analyze their relation to academic achievement (e.g., Eilam, Zeidner, & Aharon, 2009), to the best of our knowledge, a larger integration of self-regulatory constructs in the domain of educational research has not been attempted yet.

Therefore, this study aims to tackle this issue by empirically investigating a variety of self-regulatory constructs from different theoretical traditions regarding their predictive value for both academic achievement and performance in laboratory learning tasks. Specifically, by using explainable machine learning methods, we explored which constructs from the most prominent research traditions related to self-regulation in education are the most stable predictors regarding academic achievements and learning outcomes in laboratory tasks. By providing an empirical account of the relation and the overlap between constructs addressing self-regulation in education, we aim at augmenting and substantiating the existing – but mostly theoretical - analyses of this issue (Inzlicht et al., 2021) in order to inform future research programs.

2.1.1 Self-regulation in education

There are many dimensions and taxonomies by which the diverse theoretical approaches to self-regulation used in educational contexts can be potentially classified, such as their level of granularity, the core constructs involved, or the semantic similarity of the terms used to describe an approach. Depending on the scope of study, the merits of these possible categorizations may vary. In the current study, our main interest is in identifying important predictors for learning outcomes and academic achievements within the realm of self-regulation as well in analyzing their relation. Therefore, we decided to categorize different approaches to self-regulation based on the explanatory "mechanism” suggested by their core theoretical constructs. In particular, we distinguished whether their main explanatory constructs for describing how self-regulation affects learning address specific (1) activities, (2) forces, (3) dispositions, or (4) resources. Therefore, we categorized self-regulatory constructs across different research areas into learning activities (e.g., learning strategies), forces that fuel the learning process (e.g., interest and/or motivation), dispositions (e.g., personality traits), and cognitive resources (e.g., executive functioning or working-memory capacity). Of course, the allocation of some constructs and theories into these four categories may be ambiguous and debatable. For instance, Baumeister and colleagues' investigations of self-control include strong dispositional notions (e.g., trait self-control, Ent, Baumeister, & Tice, 2015) but also describe a limited resource required for self-regulation (i.e., willpower, Baumeister, Vohs, & Tice, 2007). It has to be noted, however, that these ambiguities do not necessarily reflect a flaw in our classification approach but might be rather indicative for the fact that a theoretical approach is less monolithic (i.e., more integrative) regarding its explanatory rationale. All in all, the categorization outlined below might be helpful to better understand how self-regulation is used as term for a variety of constructs that greatly differ in the way they effect learning processes and outcomes. The value of such an approach has been previously showcased in meta-analytic investigations of predictors of academic performance (Richardson, Abraham, & Bond, 2012). In order to get a more extensive picture of self-regulation in education, we decided to incorporate constructs from all four of the defined categories into our empirical study: learning activities, driving forces, cognitive resources, and self-regulatory dispositions.

2.1.1.1 Learning activities

One important area of research on self-regulation in educational contexts primarily focuses on the learning activities themselves that take place during a learning episode, such as the use of cognitive and metacognitive strategies. The central construct in this research area is referred to as self-regulated learning (SRL). Scholars investigating SRL generally aim at understanding how behavioral, cognitive, and metacognitive processes are actively regulated during learning (e.g., Azevedo, 2005, Schunk & Greene, 2018). Theoretically, these processes and their active regulation are seen as central predictors for effective learning. Although the specific focus varies depending on the underlying model used (see Panadero, 2017), metacognitive monitoring and control are the central mechanisms of regulation in SRL (Nelson &

Narens, 1994). Specifically, SRL models assume that learners are required to actively engage in multiple cognitive and metacognitive activities during learning. These activities follow a cyclical nature and typically include planning and goal setting, strategy use, monitoring, and adaption of the learning process (Greene & Azevedo, 2009; Pintrich, 2000; Winne & Hadwin, 1998, 2008; Zimmerman, 2008). A large body of empirical investigations has shown that SRL activities a – as well as interventions that aim to foster them - have a substantial positive impact on learning and achievement in laboratory settings, schools, universities and beyond (Dent & Koenka, 2016; Dignath & Büttner, 2008; Dignath, Büttner, & Langfeldt, 2008; Sitzmann & Ely, 2011; Zheng, 2016). It has become clear that the effectiveness of specific SRL activities and strategies strongly depends on multiple individual and contextual factors. For instance, in their meta-analysis Dignath and colleagues (2008) have shown that the impact of SRL on academic achievement is dependent on the specific learning strategy deployed, the academic subject under investigation, and students grade level (i.e., their developmental stage). With regard to domain specificity, laboratory studies showed similar patterns as the effectiveness of learning strategies also varied across domains (Greene et al., 2015). Therefore, to achieve desirable learning outcomes students need to engage in adequate learning activities that are adequate for the task at hand. However, studies have shown that this imposes a significant challenge as learners often rely on ineffective, shallow learning strategies (Azevedo, 2005; Azevedo et al., 2013; Narciss, Proske, & Koerndle, 2007).

In short, the central role of SRL in academic contexts is built upon extensive theoretical and empirical support. Metacognitively monitoring and controlling learning strategies in a cyclical fashion is the central component of theories and models of SRL (Panadero, 2017; Puustinen & Pulkkinen, 2001), researchers have investigated further mechanisms of SR. The activities learners engage in during learning (i.e., cognitive and metacognitive strategies) and their regulation critically determine academic success, particularly in complex or novel learning task. Specifically, students have to deploy adequate learning strategies, monitor the throughout the learning process, and make adjustments if necessary (e.g., if the currently used learning strategy was ineffective). However, research on SRL and adjacent fields of research have identified further mechanisms and processes that are essential to the regulation of learning processes, including affective and motivational processes (e.g., Efklides, 2011; Pintrich, 2003), cognitive resources (e.g., cognitive load theory, de Bruin & van Merriënboer, 2017) or self-regulatory dispositions (e.g., personality psychology Bidjerano & Dai, 2007). In the following sections we will briefly outline how these research traditions are related to leaning activities and educational outcomes.

2.1.1.2 Driving forces

Self-regulation in educational contexts is not limited to ‘cold’ cognitive processes, but also includes ‘hot’ emotional and motivational components. These *driving forces* can strongly affect if and how a learning episode unfolds and what its outcomes are. Further, emotional and motivational changes can be the outcome of learning activities, which in turn can alter further learning episodes (Boekaerts, 1996; Pintrich, 2003). The importance of motivational and emotional processes is directly or indirectly acknowledged in principle in many models focusing on *learning activities* (Boekaerts & Pekrun, 2016; Efklides, 2011; Pintrich, 2000, 2003; Zimmerman, 2000) or *self-regulatory dispositions* (cf. Inzlicht et al., 2021). Existing research typically aims at determining which and how ‘hot’ processes drive the learner to engage in and to maintain engagement in learning episodes and how these processes can be regulated. The most prominent theoretical constructs used to describe these driving forces include, among others, situation-specific value beliefs and expectancies (Eccles & Wigfield, 2002), situational interests (Hidi & Renninger, 2006), or achievement emotions (Pekrun, 2006). Recently, also the most prominent framework of emotion regulation (Gross, 2013) has been adapted to achievement and learning settings (Harley, Pekrun, Taxer, & Gross, 2019).

Empirical studies strongly support the importance of driving forces for learning and academic achievement. Meta-analyses have shown that motivation (expectancy and value) predicted school achievement across school forms and grades levels even when intelligence is controlled for (Kriegbaum, Becker & Spinath, 2018). Further, interventions that aimed at fostering student’s motivation have shown a substantial effect on academic achievement (Lazowski & Hulleman, 2016). For emotions, meta-analyses have shown that particularly negative emotions, such as boredom (Tze, Daniels, & Klassen, 2016) or math anxiety (Namkung, Peng, & Lin, 2019), have significant negative effects on learning outcomes. In addition to the detrimental effect of negative emotions on learning, beneficial effects of positive emotions have been shown in technology-based learning environments (Loderer, Pekrun, & Lester, 2020). This meta-analysis further demonstrated that emotions have a significant impact on learning processes (e.g., engagement and strategy use).

Taken together, theoretical and empirical accounts underlined the essential role of driving forces in SR and learning. These affective or motivational processes can facilitate or hinder specific learning activities, related processes, and outcomes and vice versa (Efklides, 2011; Pekrun, 2006; Taub et al., 2020). The motivational and emotional direction and energization of action provides a complementary approach to the (meta-)cognitively centered perspective of learning activities. Therefore, an integrative approach towards self-regulation, in addition to learning activities, needs to consider a variety of driving forces (e.g., motivation, interest, and emotions) and their regulation (Harley et al., 2019).

2.1.1.3 Cognitive resources

Regulating thoughts, feelings, and behaviors requires limited cognitive resources. Scholars investigating these resources usually focus on domain-general cognitive processes that are essential to exert cognitive control over behaviors or thoughts (e.g., Miyake & Friedman, 2012). The most prominently used theoretical constructs in this area are related to models of working memory structures and executive functions. In instructional research, dominant theoretical frameworks such as the cognitive load theory (Paas & Sweller, 2012) or the cognitive theory of multimedia learning (Mayer, 2014) consider working-memory resources to be the most important bottleneck for active and complex learning processes (e.g., van Merrienboer & Kirschner, 2007). Therefore, considering available working-memory resources is supposed to be highly important when designing instructional approaches, learning materials, or support measures, particularly when developing multi-media contents (Höffler, & Leutner, 2007) and other complex learning environments (van Merrienboer & Kirschner, 2007) that require learners to actively select, organize and integrate information during self-regulated learning episodes.

Theoretically, the control of working-memory resources is usually attributed to so-called executive functions, three of which are typically distinguished (e.g., Miyake et al., 2000). These functions allow to engage in the inhibition (e.g., not diverting attention to unimportant stimuli), shifting (e.g., switching between tasks), and updating (e.g., modifying representations in working memory) of working-memory contents. In additions, more recent frameworks suggest that these three components might not be at the same level but that there rather might be one more general and two more specific components of executive functioning. The general component addresses the maintenance of goal-relevant and the inhibition of goal-irrelevant information whereas the two more specific functions are required to cover shifting-specific and updating-specific aspects of cognitive control (Friedman & Miyake, 2017). Distinguishing these general and specific components seems to be more accurate in capturing the predictive value of executive functioning for different behavioral as well as clinical outcomes (e.g., substance use or procrastination behavior, Gustavson, Stallings, et al., 2017; Gustavson, Miyake, Hewitt, & Friedman, 2015).

Regardless of the specific framework or theory used to analyze the role of cognitive resources there is ample evidence for their impact on self-regulatory processes in important educational contexts (e.g., cognitive control and metacognition) as well as on learning outcomes and academic achievement (Alloway & Alloway, 2010; Cowan, 2014; Friso-van den Bos, van der Ven, Kroesbergen, & van Luit, 2013; Malanchini et al., 2019; Miyake et al., 2000). For instance, Alloway and Alloway (2010) have shown that composite measures of working memory in early stages of education are among the most powerful predictors of later academic achievement, beyond the predictive value of intelligence. Another meta-analysis showed that mathematical performance is closely related to the updating competent of working memory but that the strength of the relation between mathematical performance and working memory is dependent on the working memory measure used (Friso-van den Bos et al., 2013).

Recently studies have begun to start linking research on working memory, executive functions and cognitive to other areas of research on self-regulation in education (e.g., De Bruin & Van Merriënboer, 2017), however, these efforts remain scarce. For instance, one such study investigated multiple aspects of self-regulation, including executive functions and personal dispositions, regarding their predictive value for reading and mathematics abilities in young children (Malanchini et al., 2019). They demonstrated that latently modelled executive functions are a stronger predictor of these outcomes than other measures of self-regulation, even when controlling for fluid intelligence. This type of investigations demonstrate that cognitive resources should be considered as an essential part of self-regulation in educational contexts with substantial additive value to other constructs and concepts of self-regulation.

2.1.1.4 *Self-regulatory dispositions*

The personal dispositions that a learner brings to learning situations are the central focus of a fourth line of research related to self-regulation in education. Personal disposition, such as personality traits, are characterized by a relative stability over time and situations when compared to the constructs outlined in the previous sections, although they have been shown to develop over the lifespan (Roberts et al., 2006). Therefore, the temporal resolution of theories and models in this area of research is coarser grained and investigates the effects of self-regulation across larger parts of the 'educational lifespan'. Additionally, these theories and models typically focus on individual differences regarding general self-regulatory disposition such as trait self-control or conscientiousness (e.g., Inzlicht et al., 2020). Although trait self-control is closely associated with important life outcomes, (e.g., health, Shanahan et al., 2014), it is not necessarily predictive for in-situ self-regulation. Specifically, studies have shown that individuals with high trait self-control even tend to engage in less self-regulation on a moment-to-moment basis, because their trait self-control enables them to avoid situations with high demands on self-regulation (Hill, Nickel, & Roberts, 2014; Hofmann, Baumeister, Förster, & Vohs, 2012).In the context of education, the most commonly investigated self-regulatory dispositions include personality traits (e.g., conscientiousness; Costa & McCrae, 1998), trait self-control (Tangney, Baumeister, & Boone, 2004), grit (Duckworth, 2016; Morell et al., 2020) academic self-concept and dispositional interests (Hidi & Renninger, 2006; Marsh, 1990; Rieger et al., 2017), as well as cognitive abilities (e.g., Roberts et al., 2007).

Empirical evidence has shown that these trait-level self-regulatory constructs are closely associated with educational outcomes. For instance, reviews show that personality traits, particularly conscientiousness, predict GPA (Poropat, 2009). Similar to conscientiousness grit, which is roughly defined as perseverance and effort in the service of long-term goals, has been repeatedly associated with positive academic and life outcomes (Duckworth, 2016). Though grit was generally positively related to academic outcomes, its distinctiveness from and added value beyond conscientiousness remains debated (Credé, Tynan & Harms, 2017; Jachimowicz, Wihler, Bailey & Galinsky, 2018; Ponnock et al., 2020). Similarly, trait measures of self-control also demonstrated a significant positive relation with school outcomes (e.g., Tangney, Baumeister & Boone, 2004). These reviews and studies further have in common that the effect of general mental ability is generally controlled for, demonstrating that self-regulatory traits affect educational outcomes independently from intelligence. However, strong correlations and semantic overlap of constructs (e.g., grit and conscientiousness; Muenks et al., 2017) indicate that jangle-fallacies are probably present within this research tradition (cf. Ponnock et al., 2020).

As mentioned above (see learning activities), there have been some attempts to bridge the gap between dispositional constructs and online mechanisms of self-regulation (e.g., Bidjerano & Dai, 2007). These studies generally suggested that the effect of regulatory traits on learning might often be moderated by the use of in-situ regulatory processes.

2.1.2 Connecting constructs of self-regulation

The research approaches outlined above address issues of self-regulation in educational contexts from very different angles. Moreover, all of these research approaches have shown that 'their' self-regulatory construct has a significant impact on learning outcomes at different levels. However, the number and diversity of constructs that were introduced into research on learning and education under the umbrella term of self-regulation has also created a lack of clarity. Self-regulation research suffers from both jingle (i.e., two different constructs have the same name) and jangle fallacies (i.e., a core construct is referred to with different names). Furthermore, the varying scopes of the different research approaches, ranging from investigations of specific effects in artificial learning tasks in the laboratory to large-scale longitudinal survey studies on academic achievements leaves the question open if and which effects transfer from laboratory settings to "real" educational outcomes and vice versa. This fragmentated, intangible state of research remains a major challenge in the endeavor of identifying and fostering the most important aspects of students' self-regulation in an integrated way in order to maximize learning outcomes and academic achievements substantially and sustainably. It is the main goal of the current study to tackle this issue by investigating the predictive value of constructs from different research approaches to self-regulation across different outcome measures that range from laboratory learning tasks to school grades. More specifically, our goals are to:

- Investigate the predictive performance of self-regulatory constructs on performance in school achievement and in the laboratory learning tasks based on parsimonious, explainable models
- Identify the most robust and important features to predict learning within and across contexts
- Investigate the generalizability of these models (i.e., do the models work “outside of their context”)
- Illustrate the variance in predictions and relevant predictors
- Show connections between self-regulatory constructs

The scope of the present study is very different from most of the traditional research on self-regulation in education. Specifically, while studies from the research approaches outlined above mainly focus on the effects of a small set of key constructs and/or their interaction, we aimed at integrating the different research traditions outlined above into a larger picture. Initial studies that started do bridge the gap between two adjacent areas of self-regulation have shown that the interrelations can be additive (e.g., self-regulatory predictors with different underlying mechanisms explaining different parts of the variance in academic performance, Richardson et al., 2012), compensatory (e.g., driving forces offsetting the importance of personal dispositions, Trautwein et al., 2015) or causal sequences (e.g., learning activities mediate the effect of personal dispositions on learning outcomes, Jansen, Van Leeuwen, Janssen, Jak, & Kester, 2019). In the present study we focused on additive predictive value of self-regulatory constructs from four central research traditions on self-regulation in education (learning activities, driving forces, limited resources, and personal dispositions). To our knowledge, the present study represents the first empirical account to integrate such a broad range of measures. Therefore, we aim to provide first insights in the complementary value of self-regulatory constructs from the aforementioned research traditions. To this end, we focus on patterns of variables that predict learning best, instead of investigating the impact of individual constructs separately. Regardless of the specific approach, synthesizing different constructs related to self-regulation in education poses significant analytical and methodological challenges. For instance, variables from different research traditions greatly differ in their measurement properties (e.g., sum scores of self-reports vs reaction time data from cognitive tasks), which can negatively affect their relations to other self-regulatory constructs (Dang, King, & Inzlicht, 2020) and outcome measures (Eisenberg et al., 2019). Further, a large set of measures is required to represent the different research traditions adequately and equally. This leads to issues of multiple testing, that classical statistical approaches cannot adequately account for (e.g., Eklund, Nichols, & Knutsson, 2016). In the light of replication issues in social and psychological sciences (e.g., Ioannidis, 2005), methodological approaches that can account for these issues are required to avoid false positive findings and obtain robust, generalizable results.

One particularly promising approach to address these issues are machine learning models. This includes approaches to feature selection and engineering, i.e., the process of filtering out only the most important predictors from a large set of variables, that have been extensively researched and optimized over the last decades to identify signals even in noisy data (e.g., Saeys, Inza, & Larrañaga, 2007). In feature selection, the focus is commonly on identifying patterns rather than focusing on the role of single predictors. A potential overfitting can be avoided through the use of cross-validation procedures. More specifically, by training and testing models on different parts of the data a more robust picture of the predictive values and (more importantly) the generalizability of results is obtained. This focus on the prediction of new data in addition to the explanation of already known data is a central (but often overlooked) goal in psychological research (Yarkoni & Westfall, 2017). However, when solely prioritizing model performance (e.g., its prediction accuracy) these advantages often come at the cost of complex, black-box-like models, that allow little insights on how the predictions are made. Given that this study aims to reveal *how* constructs from different research traditions on self-regulation explain learning together, we concentrated on explainable models. More specifically, in all steps of the model development we employ machine learning method that allowed (direct) interpretation of the relation between predictors and outcome. Using such models, we optimized predictions as much as possible instead of solely focusing on prediction accuracy with no regards for interpretability.

2.2 Methods

2.2.1 Participants

Three hundred and twenty-one students were recruited for the present study. Individuals aged between 18 and 30 years with at least a university entrance qualification were eligible to participate. Further, individuals with neurological disorders and significant visual impairment (larger than 3.5 diopter – because of the quality of eye-tracking data) were excluded from this study. Lastly, only students who took part in all three sessions of the experiment were considered for analyses leading to a sample size of N = 317 (age: M = 23.26, SD = 3.02; sex: 219 female, 95 male, 3 not specified; 97.5% enrolled at a university or university of applied sciences).

2.2.2 Procedure

The present study consisted of a series of experimental tasks and surveys that were administered over three four-hour sessions. In total, 317 participants completed 25 self-report surveys and 13 cognitive tasks (e.g., working memory and executive functioning tasks) related to different facets of self-regulation in educational contexts. Moreover, five experimental learning tasks were administered to obtain measures of laboratory learning outcomes. Due to the extensive amount of testing required, data was collected in groups of up to 15 individuals in parallel. The sessions for each group were scheduled on the same day and time slot for three consecutive weeks. In few cases participations deviated from this schedule (e.g., because of sickness). The order of the surveys and tasks across all sessions was predefined to equally distribute the different types of activities (self-report surveys, cognitive tasks, and learning tasks) across sessions and to fit all self-reports and tasks into four-hour sessions including scheduled breaks half-way through each session. Experimental manipulations (e.g., different orders of texts in a reading task, using a Tablet-PC or PC for the art learning task) were randomized on participant level separately for each task that required randomization. Participants were given the option to choose between monetary compensation (32 € per session) or course credit. Furthermore, participants received a bonus when they completed all sessions (additional monetary compensation of 15 € or extra course credit).

2.2.3 Materials

All self-reports and tasks in the present study were administered digitally and displayed on 15-inch laptop screens[1]. To ensure high accuracy of the eye-tracking data recorded during the extensive sessions, longer questionnaires were split into multiple pages, so scrolling was avoided (e.g., 12 items per page for the NEO personality inventory). The selection of measures analyzed in the present paper will be briefly introduced in the next sections. A full list of all measures used in the present study can be found in Table 2.5 and Table 2.6.

2.2.3.1 Self-report surveys.

Scales and sub-scales of seven questionnaires representing different lines of research on self-regulation in education (see section 'self-regulation in education') were selected and used for analyses in the present study. An overview of these subscales and their internal consistencies are shown in Table 2.5. Missing values were treated according to the respective test manuals. If no explicit way of handling missing values was covered in the manual, scales and subscales for a participant were calculated if less than 25 % of the items of the corresponding scale were missing (rounded up to integer values, e.g., for a subscale with 3 items: 3 * 0.25 = 0.75 rounded up to 1 missing item(s) was deemed acceptable for scale calculation).

NEO personality inventory (NEO PI-R). We used the NEO PI-R (Costa & McCrae, 1998; Ostendorf & Angleitner, 2004) to capture the big five personality factors: agreeableness, conscientiousness, extraversion, neuroticism, and openness. These factors were measured on a five-point Likert scale ranging from 1: strongly disagree to 5: strongly agree. Factor scores were calculated if the participant had less than seven missing items of the 48 items in the respective facet (see Ostendorf & Angleitner, 2004). Reversed items were recoded so higher scores corresponded to higher values of the respective personality factor before sum scores were calculated.

[1] The only exception was the art learning task, where participants in one condition used 12-inch iPad® Pro's instead. In this experiment the size of the stimuli on the laptop screens was matched to the physical size of the iPads®.

Brief Self-Control Scale (BSCS). The BSCS (Bertrams & Dickhäuser, 2009) measures self-control using 13 items on a five-point Likert scale ranging from 1: not at all to 5: very much. Reversed items (i.e., negatively worded items where higher values responded to lower overall scale values) were recoded, so larger values reflect higher levels of self-control.

Cross-curricular competencies (CCC). The respective subscales from PISA assessments (Kunter et al., 2002) were used to measure specific learning strategies (control, elaboration, and memorization strategies) as well as personal and motivational variables (effort and persistence in learning, self-efficacy, and instrumental motivation). Items for these scales were captured on a four-point Likert scale ranging from 1: almost never to 4: almost always. Reversed items were recoded so higher item values responded to higher scale values.

Inventory for the Measurement of Learning Strategies in Academic Studies (LIST). To measures the use of learning strategies we used the LIST (Boerner et al., 2005, Wild & Schiefle, 1994). The questionnaire was developed based on the Motivated Strategies for Learning Questionnaire (MSLQ, Pintrich, Smith, & McKeachie, 1989). It measures a range of learning strategies (incl. cognitive, metacognitive, and effort-related strategies) with 85 items on a five-point Likert scale ranging from 1: strongly disagree to 5: strongly agree. Furthermore, the LIST includes 3 additional questions about academic achievement and time investment during university studies. After recoding higher values of each item responded to higher values of the respective learning strategy.

Long-term GRIT. The GRIT subscales consistency of interest and effort and persistence were measured using a recently developed improvement of Duckworth Quinn’s (2009) GRIT instrument, that emphasizes the long-term timescale of grit (Morell et al. 2020). The seven items of the two subscales (consistency of Interest and perseverance of effort) items were measured on a five-point Likert scale ranging from 1: not at all like me to 5: very much like me.

Emotion Regulation Questionnaire (ERQ). The emotion regulation strategies cognitive reappraisal and expressive suppression were measured through the ERQ

(Abler & Kessler, 2009; Gross & John, 2003). Over 10 items the questionnaire asks participants to indicate how they regulate positive and negative emotions and their expression on a seven-point Likert scale ranging from 1: strongly disagree to 7: strongly agree.

Domain Specific Self-Concept, Motivation, and Interest. Subject specific self-concept, interest, and motivation was assessed using a 12-item four-point Likert scale for each subject ranging from 1: not at all to 4: exactly. Specifically, for each subject self-concept was computed from a four-item, interest from a two-item, and the value of the domain from a four-item subscale.[2] After recoding higher values of each item responded to higher values of the domain-specific self-concept, interest, or value. The expectancy value was calculated as the product of self-concept and value for each domain. The mean and standard deviation of expectancy values across five domains were used as predictors for analyses.

2.2.3.2 Cognitive tasks and control measures.

In addition to questionnaire data, a broad array of psychological paradigms measuring different aspects of working memory, executive functioning and further abilities were administered. The tasks described in the following sections were implemented using Presentation® experimental software (Neurobehavorial Systems, 2018), except for the operation and reading span tasks, which were implemented in E-Prime® (Psychology Software Tools, 2018).

N-back tasks. A two-back and a three-back task were used to obtain measures of updating (Levens & Gotlib, 2012). In these tasks participants saw a series of letters and had to press a specific button when the currently displayed letter was shown N trials ago (two or three trials ago). When the letter shown did not match the letter N trials before, a different button had to be pressed. Participants completed 4 blocks with 26 items for the two-back and 27 items for the three-back task. Trials with reaction times below 200ms and outside of the range of mean reaction time ± three standard deviations for that participant were excluded. Furthermore, similar to self-report

---

[2] The remaining two items asked about the cost associated with the specific subject but were not used in the present study.

measures, we excluded participants who had more than 25 % of trials with no responses. For the present study, we computed the number of correct rejections, false alarms, hits, and misses. We used these values to calculate d’ scores according to signal detection theory (Haatveit et al., 2010).

Reading and Operation Span Tasks. Span tasks are used to measure updating/working memory capacity. In each trial of the operation span task (Turner & Engle, 1989) participants are shown a series of numbers they need to memorize. The presentation of these numbers is interleaved with arithmetical problems that participants have to solve in order to impose memory demands together with processing demands. At the end of each series of numbers, participants had to enter the memorized numbers in correct order. The reading span task (Daneman & Carpenter, 1980) is structurally equivalent, but is based on a series of letters displayed, which are interleaved with sentences that participants need to classify as true or false in order to impose processing demands during memorization. The series of items presented in a trial was enlarged over time in order to measure the maximum series length that a particular participant can handle (i.e., his or her working memory capacity) An adaptive version of both tasks was used in the present study. After three successful trials of a specific length the length of the series of numbers or letters was increased by one. We used the maximum length of series that could be correctly solved for three times as measures of working memory capacity.

Wisconsin Card Sorting Test (WCST). In the WCST participants are shown four stimulus cards with different numbers of different shapes in different colors. Their task is to match cards from two 64 card decks to these four stimulus cards according to hidden rules that they are not informed about beforehand (i.e., the cards need to be sorted either according to the types, colors or numbers of shapes). In order to allow them to derive the correct rules for matching, they receive feedback if they matched a card correctly or not. However, the rules for matching are changed constantly so that participants have to flexibly adjust to new rules as fast as possible based on the feedback they received. From this task we calculated the number of total errors made and particularly perseverance errors as a measure of how fast participants were able to switch to new rules.

Task switching paradigm. In this task participants were shown a number with a line displayed above or below it. They were instructed to indicate whether the number was smaller or larger than five when the line was displayed above the number by pressing the respective buttons. If the line was displayed below the number, on the other hand, participants had to identify whether the number was even or uneven and press the corresponding buttons. The position of the line was varied resulting in trials were the task remained the same from the previous file (non-switch trials) and trials where the task switched (switch trials). As measures of switching costs, we calculated the differences in accuracy and reaction time between switch trials and non-switch trials.

Stop signal task. Inhibition was measured using a stop signal task (Logan, 1994). In this task participants are instructed to react to a stimulus as quickly as possible, unless a stop signal is displayed (25% of trials). In this case participants were instructed to inhibit their reaction. The time interval between a target and the next stop signal is dynamically adjusted in range from 150 ms to 550 ms. More specifically, following a successful inhibition the delay between a target and the next stop signal is increased, whereas this delay is decreased following a failed inhibition. From this task we calculated the Stop Signal Reaction Time (SSRT, Matzke, Dolan, Logan, Brown, & Wagenmakers, 2013).

Stroop Tasks. A Stroop task consisting of two different blocks was used to obtain measures of inhibition. In the first block participants were shown color words in different colors and had to indicate the meaning of the word by pressing the corresponding button. This block contained incongruent trials in which the color of the word mismatched the meaning and neutral trials where the word was displayed in gray. In the second block participants were shown incongruent color words or colored x's (i.e., 'xxxxx') and had to name the color of the word/letters. From these blocks two measures of inhibition were obtained: (1) the reaction time difference in naming color words between incongruent and neutral color words (ink interference – first block) and the reaction time difference in naming colors between incongruent color words and colored letters (color word interference).

Cultural Fair Intelligence Test (CFT). General mental ability was measured using the CFT (Weiß, 2006). Participants were shown a series of figures and asked to select a figure from five options that completes the pattern or identify the option that was different from all other options depending on the block. The task consists of three blocks with 15 and one block with 11 items. Items in each block are arranged in order of increasing difficulty and each block had a time limit (four minutes for block one and two, three minutes for block three and four). For the present study, we excluded six of the 56 items from the analyses because they were potentially affected by technical difficulties in displaying all answer options correctly. Tests showed that the sum scores with and without affected items correlated highly (r = 0.99, p < 0.01), indicating that the potential issues did not affect the distribution of scores. Seven participants had no data for the last block of the task. Given that the performance of these participants in the first three block was comparable with the performance of the rest of the sample in the first three blocks, we decided to use block-level sum scores in the missing data imputation (see Missing Data Handling) before calculating the sum score across all blocks as an indicator for general mental ability.

Paper Folding Test (PFT). Participants' visuospatial ability was measured with the short version of the PFT (Ekstrom, French, Harman, & Dermen, 1976). This test comprises ten multiple-choice items, where participants have to choose the correct answer out of five options. The final item of this test was displayed incorrectly due to technical issues. Therefore, we used the sum score of the remaining nine items as an indicator of visuospatial ability. The distribution of completed items indicated that only few participants reached the affected item, indicating that discriminatory power of the test was largely maintained.

Reading comprehension test (LGVT* 5-12+*). Reading ability was assessed using the LGVT 5-12+ Schneider, Schlagmüller, & Ennemoser, 2017). In this test participants can spend four minutes to read a text with 47 gaps and to fill in the correct words (out of three answer options) for as many gaps as possible. For this study we adapted the paper-pencil based test for a computer-based use. Specifically, the gaps in the text were implemented as drop down selection menus. When clicking on a gap, participants could select the answer from a drop-down menu. Furthermore, the text was split into parts that fit on a screen without scrolling. Forward and backward buttons were used to navigate to the next or previous part of the text. Reading comprehension and accuracy were calculated according to the test manual (Schneider et al., 2017).

2.2.3.3 Academic performance.

Participants were asked to indicate their last grades for specific subjects at the end of grammar school, their overall final grade in grammar school, as well as their current GPA in several questionnaires (e.g., LIST, Boerner et al., 2005). We used self-reported grades at the end of grammar school as measures for academic performance for several reasons. First, the self-reported grades in arts, biology, history, math, and physics map directly onto the content domains of the laboratory tasks in this study. Second, these grades are directly comparable between participants as they measured academic performance in these subject at the same timepoint in their education on roughly the same level. The current GPA in their studies on the other hand is hard to compare across semesters and academic disciplines. The self-reported grades were administered as part of the demographic questions at the start of the first experimental session before any other task or self-report were administered. Grades were re-scaled ranging from 0 to 5 so higher values corresponded to higher academic achievement.

2.2.3.4 Learning outcomes in laboratory learning task.

Five experimental learning tasks and environments were developed or selected from previous research in order to measure specific and objective learning outcomes in different content domains (i.e., arts, biology, history, math, physics). These learning tasks and environments were designed to be representative of typical learning tasks and materials in the respective content domains. For all learning tasks that contained experimental manipulations, we controlled if the scores used for analyses significantly differed between experimental conditions (see Preliminary Analyses for details).

Art-history learning task. In the art learning task participant learned about five core topics of art-history either by using a PC or Tablet-PC. Specifically, we developed a multi-perspective hypermedia learning environment with art-historic contents based on cognitive flexibility theory (Jacobsen & Spiro, 1995) and on our own research on the role of executive functions in multi-perspective hypermedia learning environments (Kornmann et al., 2016). Before learning with the learning environment, participants completed a 30-item pre-test measuring art-history knowledge relevant to the contents of the learning environment. In the subsequent learning phase participants were tasked to learn as much as possible about the 20 artworks featured in the learning environment by freely exploring the contents and were provided with exemplary questions to potentially guide their learning. For the present study we included the percent of correctly answered questions in the 30-item posttest that followed the learning phase as a learning-outcome measure.

Biology learning task. The learning task covering biological materials consisted of six expository texts about exotic species. All of these texts contained global (between paragraphs) and local contradictions (between neighboring sentences). Before reading the texts participants prior knowledge was measured with multiple-choice questions. Participants were instructed to learn as much as possible about the species. The order of texts was randomized. After participants finished reading a text, they had to answer 11 multiple-choice questions before moving to the next text. As a performance measures we calculated the percent correct across all texts for factual knowledge items (three per text), inference items (three per text), items addressing local contradictions (two per text) and items addressing global contradictions (two per text). These scores were averaged to obtain the overall learning-outcome measure.

History learning task. The history learning task consisted of a hypertext environment about the history of the Panama Canal containing hyperlinks that were relevant or irrelevant to the overall task. The goal was to learn as much as possible about the history of the Panama Canal. Depending on the experimental condition, participants received different multiple-choice questions prior to the learning phase, that either (or not) covered contents from irrelevant links embedded it the environment. Subsequent to learning with the hypertext, participants completed a post-test containing ten inference questions for verification and 24 multiple-choice questions addressing contents from hyperlinks relevant and irrelevant for the overall learning goal. For the present study we used the average score of the inference questions and those eight multiple choice items that addressed contents from the hyperlinks relevant to the learning goal as learning-outcome measure.

Math learning task. For the math learning task, we used *HyperComb* (Gerjets, Scheiter, & Catrambone, 2006), a hypermedia learning environment focused on the topic of calculating the probability of complex events. In this environment students completed a brief introduction to the topic including a declarative knowledge test, before they engaged in an example-based learning phase. Specifically, they were instructed to learn about four problem categories using worked-out examples with differing levels of elaboration. Finally, as a post test, participants worked on a math exam consisting of mathematical problems and declarative knowledge questions. We computed the proportion of correct answers to five isomorphic and six transfer problems from the posttest as learning-outcome measure for analyses.

Physics learning task. An expository text about the moon from OECD's Programme for International Student Assessment (PISA, Kunter et al., 2002) and the corresponding test items were used as a physics learning task. For our analyses we used the percentage of correctly answered items from the six multiple-choice retention questions after the text as laboratory learning-outcome measure in physics.

2.2.4 Analytical procedure.

Prior to model development, we employed a feature engineering pipeline that consisted of several preprocessing steps. Specifically, the analyses in the present study included the generation of groups for classification, missing data handling, a multi-step modelling procedure (incl. hyperparameter selection and selection of model performance measures), that will be outlined in the following sections.

2.2.4.1 Generating groups of participants for classification models

Participants were split into top and bottom 30% according to (1) their average self-reported grades across the five domains as well as according to (2) their average learning outcomes in the five laboratory learning tasks. For this purpose, all outcome measures were rescaled to scores ranging from zero to one. For laboratory learning tasks, the average score across all learning outcome measures within a task (e.g., the isomorphic and transfer problems in the math task) was calculated to ensure that each domain carried the same weight for the overall learning outcomes. Then for grades as well as for the calculated scores for laboratory learning outcomes participants were grouped based on their average scores across the five domains unless they had more than two missing values in the respective scores. Specifically, participants with average grades or laboratory scores below the 30th percentile were assigned to the bottom (low performing) group and participants above 100 - 30 = 70th percentile were assigned to the top (high performing) group. The classification algorithms were trained to predict if a participant belonged to the top or bottom group. We further varied cut-offs in one percent steps between top and bottom 20% and top and bottom 40% to assess (1) how well more or less extreme groups regarding academic performance or laboratory learning outcomes can be differentiated and (2) how the importance of specific features is depending on cut-offs (e.g., some variables might only yield predictive value when more extreme groups are compared).

2.2.4.2 Missing Data Handling

To replace missing values in our predictors, we used multivariate imputations by chained equations (MICE, White, Royston, & Wood, 2011). In detail, for each predictor with missing values, a Bayesian ridge regression predicting the variable selected for imputation with all remaining predictors was fit on non-missing values and used then to predict missing values. This procedure was repeated for each predictor and carried out in ascending order of the missing values per variable (i.e., the predictor with the least amount of missing values was imputed first, the one with most missing values last). Additionally, for categorical variables or variables with values limited to full integers, we rounded the missing values to full integers (e.g., if the imputed operation span of a participant was estimated to be 5.1, 5 was used as the imputed value instead). The outcome variables, i.e., grades or performance in laboratory tasks, were not imputed.

Modelling Approach

Our main goal was to identify the most stable predictors for school grades and laboratory learning outcomes using parsimonious and explainable machine learning models. To this end we evaluated multiple modelling approaches. Specifically, multiple feature selection procedures (e.g., step-wise-forward selection, recursive feature elimination, and model-based approaches) and a selection of classification algorithms (i.e., linear support vector machines [SVMs], AdaBoost classifier, bagging classifier, and random forest classifier) were employed. With regard to the feature selection method, we decided to choose a modelling approach that selected a sparse and consistent range of features throughout the cross-validation. For the classification model, we tested whether more complex modelling approaches with indicators of feature importance (i.e., permutation importance) yielded more accurate predictions than linear SVMs. These tests revealed that neither the AdaBoosting nor Bagging Classifier (based on linear SVMs) were significantly more accurate than linear SVMs and therefore, they were disregarded.

The final feature selection procedure in the present study was a two-step approach. (1) A univariate 'pre-selection' of features was conducted through false discovery rate (FDR) corrected p-values using the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995) on analyses of variance (ANOVAS) comparing mean values for each predictor between top and bottom n % of participants. (2) A 10-fold cross-validated recursive feature elimination (RFECV) was used to limit the features used for predictions to include only necessary features. Specifically, we used a logistic regression with L1-regularization as basis for this step. A L1-norm was deemed particularly suitable for this step because it generally produces sparse feature sets. The area under the receiver operating characteristics curve was used as the performance criteria for the RFECV. Using the features selected through this approach, we then used linear SVM and a random forest classifier to predict the group membership of a participant.

To avoid overfitting and to increase the generalizability of our predictions the steps outlined above were performed with a ten-fold stratified cross-validation. In this approach, the data set is divided into ten equal parts (folds). One of these folds is then selected as test dataset while the remaining nine folds are used as training dataset. For our modelling approach, the feature selection was carried out and the training of the classification algorithm was fit on the training dataset. For the RFECV, this training data was further split into ten folds using an additional ten-fold stratified cross-validation on the training data set. The fitted prediction model is then used to predict the test dataset. This procedure was repeated ten times, with a different fold used for testing purposes each time. The average performance and variance of models across the ten folds are used as outcomes.

We decided to incorporate two more steps to guarantee as much stability and generalizability as possible. First, we decided to systematically vary the cutoff that defines the top and bottom performing students in school and in laboratory tasks.

Specifically, in one percent steps between 20% and 40% we repeatably split the participants into top and bottom performer for grades and laboratory learning outcomes separately. This approach allowed us to estimate how well participants performing at certain levels can be differentiated and how important predictors are depending on the cutoff. Second, we repeated the cross-validation procedure with randomly shuffled data 100 times for each cut-off to get a more comprehensive picture of the variance and stability of our predictions in our sample and to identify niche predictors that might not be consistently selected but be relevant in specific subsamples. This approach yielded an ensemble prediction averaged across the 100 runs for each cutoff. Lastly, the final prediction model (SVM or random forest classifier) was based on the average performance across the hundred runs for each cutoff. For instance, at the 40% cutoff predicting grades with SVM was more accurate on average than predictions with random forest classifier. Therefore, SVM were used as classification algorithm for grades at this cutoff.

2.2.4.3 Hyper-parameter selection

Hyper-parameter were tested in preliminary tests using grid searches in 10x10-fold nested cross validations. In the feature selection we tested different p-value cutoffs for the pre-selection step (i.e., 0.05, 0.01, and 0.001). The preliminary analyses showed that models predicting laboratory learning outcomes were most accurate with the default p-value of 0.05 while the models for grades worked best with a stricter corrected FDR p-value of 0.01. We further examined different cost parameters for the SVM (ranging from .001 to 10). These preliminary tests showed that in runs where more feature were selected, lower cost parameters were more optimal. Accordingly, we decided to use the inverse of the number of features as cost parameter for all models. The random forest classifiers were computed with 100 trees, Gini impurity to measure the quality of splits, and no max depth. Lastly, comparisons of models with these fixed hyper-parameter and models with varying hyper-parameter using grid search showed no advantages in accuracy for varying hyper-parameter, which substantiated the validity of the selected hyper-parameter.

2.2.4.4 Model performance parameters

Accuracy. The proportion of correctly predicted low and high grades and laboratory learning outcomes in the test sample was averaged across all folds of each run. Means and standard deviations for the group sizes were calculated over all runs and used as a stable indicator of model performance and its variation.

Selected Features. As an indicator if and how consistent a predictor is used to classify grades and laboratory learning outcomes, we recorded the number for times a predictor was selected by the two-step feature selection procedure and used for prediction across all folds of each run (see section Modelling Approach). We then used the percentage of models in which the feature was included as an indicator of how consistently the predictors was selected for prediction. In addition to identifying the most consistent features for prediction this approach also displayed the variety of features that show significant differences between high and low grades and laboratory learning outcomes across different group sizes and sub-samples of our sample (i.e., random splits for the cross validation).

Permutation Importance. We used permutation importance tests to analyze how important selected features were for the models. In this procedure, a feature in the test data set is replaced by a random permutation of its own values. This guarantees that the distribution of values in this variable is maintained, but the values are randomly shuffled between participants. Then the difference in prediction accuracy between the model using the regular data set and the shuffled data set is recorded. This process is repeated ten times (with different shuffling patterns) for each variable. In the present study we used the average change in accuracy across the ten repetitions for each feature as a measure of importance.

Generalizability to the other outcome. To test if our models can meaningfully predict the opposing outcome (i.e., grades models predicting laboratory learning outcomes and vice versa), we used the fitted models for grades and laboratory learning outcomes to predict the other outcome for the same group size as the original model for all participants that were not included in the training data set of the current fold. Specifically, this included all participants that did not belong to the top or bottom n% of performers for the original outcome but were in either group for the other outcome as well as all participants in the test data set that belonged to the top or bottom performers for the other outcome. Values for these predictions were computed in the same way as the model accuracy (see section ‘accuracy’).

2.2.4.5 *Control variables*

To test if some of the effects outlined above were just caused by underlying differences in general mental ability and/or reading ability we reran the analyses outlined above with general mental ability (Weiß, 2006), visuospatial ability (Ekstrom et al., 1976), and two measures of reading ability (Schneider et al., 2017). Additionally, we further included academic self-efficacy (Kunter et al., 2002) in these models. Initially, self-efficacy was considered as a predictor, but not included in the reported models because the wording of the items capturing self-efficacy were too closely linked to performance in school and grades. Together with the retrospective self-reported nature of the school outcomes we used, this phrasing potentially enhanced potential causal issues (i.e., self-efficacy indirectly covering parts of retrospective performance in school, which is one of our outcome measures). These models will be briefly reported and compared to the 'main' models to show the potential effect of these control variables (or lack thereof).

2.3 Results

2.3.1 Preliminary analyses.

2.3.1.1 *Missing data imputation*

Investigations of missing data revealed that only four participants and two variables (the reading and operations span tasks – due to a technical issue in the first days of data collection) had more than 10% missing data. To test the effect of the MICE imputation, we first descriptively compared variables before and after imputation. These descriptive observations showed that none of the mean values changed by one percent or more after imputation. However, three variables showed a change in standard deviation that exceeded one precent (reading span 8%, operation span 10%, stop signal reaction time 3%). Given that our analytic approach required complete data and an exclusion of missing data led to a significant decrease in sample size (n = 232 without missing values in predictors) we only tested how the exclusion of the most severe occurrences of missing values affected our results. Specifically, we re-analyzed the data by testing how excluding participants and variables with more than

10% missing values affected our results. These alternative analyses showed an identical pattern of results to the findings outlined below, indicating that the imputation did not affect the outcomes of the present study. Thus, in the remainder of the paper, we report all data with MICE imputation.

2.3.1.2 Outcomes measures

To test the validity and distinctiveness of the outcome measures used in the following analyses, correlations between the average grades across the five domains, the average learning outcomes across the five laboratory learning tasks, and a self-reported final grade at the end of grammar school were calculated. Results showed (1) that the mean grades across the five domains represented the final grade at the end of grammar school well (r = -0.81, p < .001), (2) that grades and laboratory learning outcomes were correlated but still showed sufficient differences (r = 0.37, p < .001), and (3) that the relation between grades and laboratory learning outcomes was lower for the average performance across five domains (r = 0.37, p < .001) than the final grade at the end of grammar school (r = -0.45, p < .001). This indicated that the average grades across the five domains were a suitable outcome measures that closely corresponded to other measures of academic performance in school. Further, the average performance in laboratory learning task representing the same academic domains, was related to performance in school but also showed sufficient differences (i.e., ca. 14% shared variance between grades and laboratory task performance). This further indicated that individuals with high (and low) grades not necessarily also showed high (and low) performance in laboratory learning tasks.

2.3.2 Prediction accuracy

Overall, across group sizes ranging from top vs. bottom 20% to 40%, the ensemble predictions classified participants from the high and low grades groups with 79.43% accuracy (SD = 3.91%) and participants from the high and low laboratory learning outcome groups with 74.66% accuracy (SD = 3.39%). In these predictions the models for grades selected 3.63 features on average (SD = 0.71) while models for laboratory learning outcomes selected 8.33 features (SD = 0.94). We further tested if the obtained distribution of accuracies was different from a distribution of random predictions (i.e., the same models predicting randomly shuffled outcomes) through permutation tests (Odén & Wedel, 1975). Results showed that the ensemble predictions for grades (random prediction accuracy: 50.04%, $p < .001$) as well as laboratory learning outcomes (random prediction accuracy: 50.16%, $p < .001$) were significantly more accurate than random predictions.

The range of average accuracies across all runs for each group size was 73.92% (40% group size) to 85.57% (20% group size) when predicting grades and 70.08 % (40% group size) to 80.12% (20% group size) when predicting laboratory learning outcomes (see Figure 2.1). Separate one-way ANOVAS further showed, that models predicting grades were significantly more accurate than models predicting laboratory learning outcomes (all: $Fs(1, 99) \geq 25.13$, $ps < .001$).

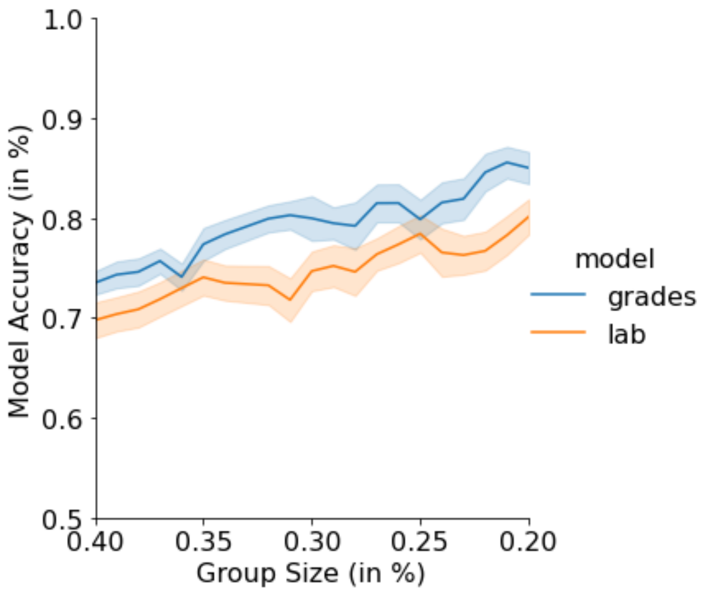


Figure 2.1. *Prediction accuracies by group size*

Correlational analyses were conducted to test if the group size was related to model accuracy and showed that ensemble predictions for grades and laboratory (r = -.857, $p < .001$) and laboratory learning outcomes (r = -.770, $p < .001$) were significantly more accurate with more extreme groups for comparison (see Figure 2.1). This indicated that differences in self-regulatory variables between very high and very low performing students were more pronounced than between less extreme groups.

We further tested if the number of features selected was related to model accuracy. Correlations revealed that models predicting grades were significantly more accurate if more features were selected (r = .382, $p < .001$), whereas for models predicting laboratory learning outcomes prediction accuracy significantly decreased with increasing number of features ($r = -.233$, $p < .001$). However, as outlined above, the laboratory learning outcome models generally selected more features than the grade models on average.

2.3.3 Selected Features

The ten most consistently selected predictors in both ensemble predictions are displayed in Table 2.1 and Table 2.2 (for mean values by performance groups see Table 2.7 and Table 2.8). Overall, the majority of features was included in a model at least once. Only elaboration strategies (Kunter et al., 2002), temporal switching cost, and the stop signal reaction time were not included at any point.

For predictions of the top and low groups regarding grades, regardless of the group size selected, the mean motivation (i.e., expectancy value) across the five content domains investigated in the present study was included in all models. Additionally, invested effort and working memory capacity (d' in 3-back task) were relatively consistently selected (> 70%) depending on the range of group sizes selected. The remaining features among the ten most commonly selected predictors were situationally to rarely included in predictions (see Table 2.1).

In models for laboratory learning outcomes three highly consistent features were identified. Rehearsal strategies, working memory capacity (d' in 3-back task) and motivation (mean expectancy value) were included in the vast majority of models (> 90%) across all group sizes. Furthermore, openness, switching accuracy, reading span, d' in the 2-back task and time management were relatively consistently selected (> 70%) in at least one of the group size ranges (> 30% group size or ≤ 30% group size, see Table 2.2).

2.3.4 Feature Importance

To assess the contribution of the selected predictors to each model we calculated the permutation distribution, which provides an estimate how much accuracy is lost if the feature is replaced by random values. These investigations showed that models predicting grades were largely driven by the mean expectancy value with an average loss of accuracy by 24.39%. In parallel with the frequency of inclusion in models, only effort (2.12%) and working memory capacity (1.06%) achieved a consistently positive contribution to the models. Lastly, the importance of all top features was larger with smaller group sizes when compared to larger group sizes (see Table 2.1).

Investigations of permutation importance in models of laboratory learning outcomes painted a more diverse picture. First, working memory measures showed the largest impact on model accuracy when included in the model (i.e., 3-back d': 6.31%, 2-back d' 3.07%), but were not the most frequently selected features (see Table 2.2). Further, the overall contribution of single features was much less pronounced, and the permutation importance values were more evenly distributed (i.e., all of the consistently selected features also showed positive contributions between 0.86% and 6.31% to the model on average). Lastly, the importance relative to the ranges of group sizes showed distinct patterns. Specifically, while the 3-back d', rehearsal strategies, openness, accuracy switching costs, and operation span showed a larger importance in models with smaller group sizes, the 2-back d' and mean expectancy value showed an opposing pattern, where its impact on model accuracy was more pronounced for larger group sizes (see Table 2.2).

Table 2.1

Top 10 most consistently selected features in grade models

Predictor	Overall		> 30% group size		≤ 30% group size	
	N_{sel}	Importance	N_{sel}	Importance	N_{sel}	Importance
Mean Expectancy Value	100.00	24.39	100.00	23.74	100.00	24.91
Effort (LIST)	77.61	2.12	67.86	0.98	90.18	3.05
D Prime (3-back)	45.01	1.06	37.96	0.31	70.51	1.67
D Prime (2-back)	26.46	-0.04	48.07	-0.11	17.29	0.01
Control Strategies	26.46	-0.16	10.26	-0.18	50.36	-0.15
Effort and Persistence in Learning	18.23	-0.06	21.10	-0.24	14.13	0.08
Reading Span	16.60	0.04	5.63	-0.06	33.08	0.12
Mean Interest	11.91	0.07	16.11	-0.03	6.84	0.16
Conscientiousness	6.28	-0.03	10.67	-0.05	3.95	-0.01
Self-Control (BSCS)	2.36	-0.02	3.86	-0.05	0.94	0.00

Note. All values in percent. N_{sel} = Selected in percent of models, Importance: permutation importance

Table 2.2

Top 10 most consistently selected features in models for laboratory learning outcomes

Predictor	Overall		> 30% group size		≤ 30% group size	
	N_{sel}	Importance	N_{sel}	Importance	N_{sel}	Importance
Rehearsal strategies (LIST)	99.98	2.57	99.98	2.10	99.98	2.95
D Prime (3-back)	99.09	6.31	99.26	4.82	99.43	7.52
Mean Expectancy Value	94.93	2.52	92.38	2.61	95.16	2.44
Openness	84.47	2.16	92.49	1.77	73.75	2.48
Switching Accuracy Difference	78.74	1.52	70.32	0.79	93.17	2.11
Reading Span	74.88	0.86	54.51	-0.01	95.67	1.57
D Prime (2-back)	71.06	3.07	87.00	4.87	46.34	1.59
Time Management (LIST)	67.88	0.32	50.19	-0.07	78.56	0.65
Memorization Strategies (CCC)	52.32	0.06	64.58	0.44	27.54	-0.25
Mean Interest	32.98	0.08	56.41	0.21	8.25	-0.02

Note. All values in percent. N_{sel} = Selected in percent of models, Importance: permutation importance in percent

2.3.4.1 Generalizability of models

To test if the models we used were specific to one outcome (i.e., grades or laboratory learning outcomes) or worked regardless of the outcome measure, we investigated how well a model trained on grades predicted laboratory learning outcomes and vice versa. Results showed that both models for grades (Accuracy: M = 53.69%, SD = 2.16%) and laboratory learning outcomes (Accuracy: M = 65.17%, SD = 2.07) performed worse when predicting the other outcome than on the outcome they were built for. Specifically, permutation tests showed that models trained on grades predicted laboratory learning task performance significantly less accurate than they predicted grades (p < .001) and models trained on laboratory learning task performance predicted grades with significantly lower accuracy than the outcome they were trained on (p < .001).

2.3.4.2 Effects of control variables

Models including the five control variables (general mental ability, visuospatial ability, academic self-efficacy, and two measures of reading ability) in addition to the predictors included in the models above yielded an average accuracy of 79.83% (SD = 4.37%) for grades and 76.78% (SD = 4.21%) for laboratory learning outcomes. In these predictions the models predicting grades selected 5.50 features on average (SD

= 0.96) while models predicting laboratory learning outcomes selected 10.04 features (SD = 1.35). Permutation tests showed that on average these models were slightly but significantly more accurate than models without control variables for grades (79.83% vs. 79.43%, $p < .01$) and for laboratory learning outcomes (76.78% vs.74.66%, $p < 0.001$). Follow-up permutation tests, broken down into larger (> 30%) and smaller group sizes (≤ 30%) showed that models predicting grades where significantly more accurate with control variables for larger group sizes ($p < .001$) but not for smaller group sizes ($p = .329$), whereas the prediction of laboratory learning outcomes models with control variables were significantly more accurate for both ranges of group sizes (both $ps < .001$).

With regard to selected features, three control variables were consistently included in models predicting grades (see Table 2.3) and laboratory learning outcomes (see Table 2.4). Specifically, general mental ability, visuospatial ability and academic self-efficacy were among the ten most commonly selected features for both outcomes. Reading abilities did not seem to play an additional role for predicting learning outcomes. Importantly, none of the dominant features in the models without control variables were removed when including the control variables.

Table 2.3

Top 10 most consistently selected features in grade models with control variables

Predictor	With control		Compared to no control	
	N_{sel}	Importance	Δ_N	$\Delta_{Importance}$
Mean Expectancy Value	100.00	17.13	+0.00	-7.26
Self-Efficacy	91.97	2.95	-	-
Effort (LIST)	79.62	1.95	+2.01	-0.17
General Mental Ability	91.97	-0.05	-	-
D Prime (3-back)	46.79	1.76	+1.78	+0.70
Visuospatial Ability	35.96	0.16	-	-
Control Strategies	22.39	0.04	-4.06	+0.20
Reading Span	20.23	0.72	+3.63	+0.68
D Prime (2-back)	19.72	-0.05	-6.74	-0.01
Effort and Persistence in Learning	11.94	-0.39	-6.29	-0.33

Note. All values in percent. N_{sel} = Selected in percent of models, Importance: permutation importance

Table 2.4

Top 10 most consistently selected features in learning outcome models for laboratory learning outcomes with control variables

Predictor	With control		Compared to no control	
	N_{sel}	Importance	Δ_N	$\Delta_{Importance}$
Rehearsal strategies (LIST)	99.54	2.59	-0.44	+0.02
General Mental Ability	99.01	3.83	-	-
Visuospatial Ability	98.89	2.54	-	-
Switching	84.88	2.10	+6.14	+0.58
Openness	78.80	1.47	-5.67	-0.69
Mean Expectancy Value	73.14	1.00	-21.79	-1.52
Self-efficacy	70.20	1.54	-	-
D Prime (2-back)	62.96	1.48	-8.10	-1.59
D Prime (3-back)	59.78	1.12	-39.31	-5.19
Memorization Strategies (CCC)	52.74	0.48	-0.42	+0.42

Note. All values in percent. N_{sel} = Selected in percent of models, Importance: permutation importance

2.4 Discussion

Our machine-learning approach provided several insights into the predictive value of self-regulatory constructs for academic outcomes and learning outcomes in laboratory learning experiments. Most importantly, our results showed that key variables from all four introduced research traditions related to self-regulation in education (learning activities, driving forces, limited resources, and personal dispositions, see Introduction) showed substantial predictive value for both outcome measures. In particular, the ten most consistently selected predictors in both ensemble predictions comprised variables describing (1) learning activities such as different strategies, (2) driving forces such as interest or expectancy-value, (3) self-regulatory dispositions such as openness or conscientiousness, and (4) cognitive resources such as working-memory capacity of executive functions. Therefore, the different types of explanatory rationale that are reflected by these theoretical concepts seem all to play an important role in covering the overall landscape of self-regulation in educational contexts. Thus, there seems to be no single best explanatory "mechanism” for self-regulation in terms of these four types of core theoretical constructs. Rather, these constructs might complement each other regarding their predictive role for learning success. These results were not only found at the theoretical but also at the methodological level: While the specific constructs that were consistently selected for prediction and their importance differed depending on the outcome, we did not find that a particular type of measurement approach (i.e., surveys or behavioral tasks) showed a general lack of predictive power (in contrast to Eisenberg et al., 2019). Rather, our study yielded that (coarser grained) survey data and (finer grained) behavioral measures in conjunction provided the best predictive value for important educational outcomes (i.e., grades at the end of grammar school) as well as for the performance in more restricted laboratory learning tasks. Together, these findings highlight that further attempts to integrate theories and studies of self-regulation in education should consider constructs across different theoretical and methodological approaches that may differ regarding their explanatory rationale as well as regarding their levels of granularity. Interestingly, previous studies on self-regulation outside of educational contexts have painted a different picture by repeatedly finding a disconnect between survey data, behavioral tasks, and real-life outcomes which was attributed to jingle fallacies and measurement issues (i.e., Dang et al., 2020; Eisenberg et al., 2019).

Overall, our modelling approach showed that grades and laboratory learning outcomes can be predicted accurately across different group sizes with parsimonious and explainable models. As expected, more extreme groups (i.e., top vs. bottom 20%) were predicted more accurately, which indicated more pronounced differences in self-regulation than in less extreme comparisons (e.g., top vs. bottom 40%). Although our models showed that both, grades and laboratory learning outcomes can be classified with good accuracy using a small set of predictors, it could be demonstrated that predictions of grades were significantly more accurate than predictions of laboratory learning outcomes, regardless of the group size selected. A potential explanation lies in the more complex situations that school settings present when compared to laboratory learning tasks. Here students have to repeatedly initiate and maintain learning efforts in different contexts across extended periods of time (e.g., homework in addition to learning in classes, Trautwein, 2007), which may result in more pronounced accumulated differences between high and low performers. Taken together, our results showed that students with superior grades or laboratory learning outcomes demonstrated significantly different levels of self-regulation prerequisites than their counterparts with inferior grades or laboratory learning outcomes. These findings are well in line with the numerous investigations linking different aspects of self-regulation to educational and other life outcomes (Baumeister et al., 1994; Burnette et al., 2013; Dignath et al., 2008; Moffitt et al., 2011; Robson et al., 2020; van Genugten et al., 2017).

With regard to the most important predictors our results indicated that performance in school and laboratory learning tasks have shared as well as distinct self-regulatory demands and prerequisites. We found that the most consistent predictors were a measure of motivation (expectancy-value) and a working-memory measure (3-back d' prime) as these two measures were the only features that were included in most models for both outcomes. The importance of these features for the prediction, however, showed an opposing pattern depending on the outcome: Whereas predictions of grades were heavily dependent on motivation, classifications of laboratory learning outcomes were more reliant on the working memory measures. This indicated that successful self-regulation across educational contexts has shared underlying mechanisms that jointly contribute to successful learning. In line with previous research, we found that motivation (Kriegbaum, Becker, & Spinath, 2018) and working memory capacity (e.g., Alloway & Alloway, 2010) are essential self-regulatory factors for educational success. Our study extends these finding by showing that these self-regulatory constructs can demonstrate their large importance even when numerous other constructs are considered simultaneously. Furthermore, our models suggest that both constructs are jointly required to succeed, for instance, because the advantages of greater cognitive abilities can only translate to better outcomes when students are motivated to engage with the material and vice versa (e.g., Zimmerman, 2000).

Models for both outcomes consistently included predictors from all four research areas referred to in the design of this study (i.e., research on learning activities, driving forces, limited cognitive resources, and self-regulatory dispositions, see Introduction). In addition to motivation as a driving forces and working memory as cognitive resource the most important predictors for both outcomes included learning activities and personal dispositions. In models predicting grades, effort (as a personal disposition) and to a lesser extend control strategies (as a learning activity) where frequently used. For laboratory learning outcomes on the other hand, openness (as a personal disposition) and (avoiding) rehearsal strategies (as a negative learning activity, see Table 2.8) were among the most relevant predictors. Further important predictors for grades primarily included personal dispositions whereas for laboratory learning outcomes cognitive resources and learning strategies were most common included in the models. These findings directly relate to previous research in multiple areas. Theories of SRL and studies on the relation between personality, SRL strategies and learning outcomes suggest that SRL processes moderate the effect of personal dispositions on learning outcomes (Bidjerano & Daj, 2007; Chamorro-Premuzic, & Furnham, 2003; Schunk & Greene, 2018). In this context, the more pronounced importance of trait self-regulation measures (e.g., conscientiousness, self-control) in school contexts as compared to laboratory settings is in line with previous research as studies have shown that individuals with higher trait self-regulation are often able to circumvent real-life situations that require high levels of state self-regulation and even show lower levels of state self-regulation (Inzlicht et al., 2021; Hill, Nickel, & Roberts, 2014; Hofmann, Baumeister, Förster, & Vohs, 2012). While this can be effective in schools where long-term preparation is possible, in laboratory settings such behavior is not applicable because as the tasks at hand can't be avoided (other than not participating) and often require high in situ self-regulation by design. As studies on SRL have shown, for successful learning overreliance on surface-level learning strategies is detrimental when learning about complex topics (e.g., Azevedo, 2005; Azevedo et al., 2013; Narciss et al., 2007). In line with these findings our models show that surface level strategy use (e.g., rehearsal and memorization) was a more prominent negative predictor in laboratory settings.

These differences between predictions of grades and laboratory learning outcomes were further underlined by a lack of generalization of both models to the other outcome. However, we did not find a lack of predictive value of cognitive task measures (e.g., working memory and EF) for life outcomes demonstrated in an integrative study on SR in clinical settings (Eisenberg et al., 2019). Instead, our findings suggest that learning in the laboratory functions differently and/or measures different aspects than learning in school. To ensure more transfer between research findings and educational practice – a central goal of educational research – more systematic investigations of learning in both settings are required. Our analyses further showed that the contributions outlined above are relatively independent of mental ability, reading comprehension and other more general control measures. In particular, including measures of general mental ability, visuospatial ability, reading comprehension or academic self-efficacy only led to slight increases in prediction accuracy. More importantly, even though the frequency of selection and importance of some predictors decreased when including these control variables, particularly for working memory measures, none of the features discussed above was rendered completely obsolete through the inclusion of control variables. In line with previous studies this showed that different self-regulatory constructs explain learning success independent of the effects of intelligence (e.g., Alloway & Alloway, 2010; Chamorro-Premuzic, & Furnham, 2008).

However, the present study also faced some limitations and challenges that need to be considered when interpreting the results at hand. One such issue involves the chronological order for measuring academic achievement (i.e., grades). Given that our sample was rather diverse in their field of study and study progress, we opted to compare them by their grades in grammar school, which are comparable for all participants. However, this introduced the issue that the academic performance was achieved before the study took place and thus could only be measured through retrospective self-reports. Therefore, we cannot exclude reverse causal relation between the predictors and grades (e.g., higher levels of self-efficacy caused by higher academic achievement, Marsh & Martin, 2011). To account for this issue as much as possible within our study design and analyses, grades were measures at the very start of the first experimental session to rule out further effects of other surveys on self-reported grades. Furthermore, we excluded questionnaire data as predictors from the analyses that were to closely be linked to grades or to performance relative to other students in school (e.g., the academic self-efficacy questionnaire). In further research, long-term longitudinal studies combining survey and behavioral measures will be needed to unravel the development and causal interconnection of the manifold processes included in self-regulation in greater detail.

Further our scope with regard to the investigation of learning activities and cognitive resources was limited to 'trait-like' accounts of these constructs. For instance, we used questionnaires to measure SRL strategies as specific types of learning activities, which is still debated with a large body of research suggesting that such processes might better be measured using trace data (e.g., Winne & Perry, 2000). Similarly, we measured different aspects of working-memory capacity and executive functioning only single session and used these data for predictions of laboratory learning outcomes that were potentially obtained in a different week (if the corresponding learning task was part of a different session). In doing so, we assume a certain stability of these measures over time, which is another issue debated in the literature (Eisenberg et al., 2019). However, whereas an in-situ measurement of executive functioning and working-memory capacity as well as process measures of learning strategies would certainly provide a great added value, they would have not been feasible in a study with this scope and they would have been unobtainable with regard to the school-related outcomes. We alleviated this issue by administering multiple measures for the different aspects of cognitive functioning as well as for learning strategies spread over the three experimental sessions and by using well-established and consistent performance measures. Accordingly, our results showed that both aspects of self-regulation (learning activities and cognitive resources) were well represented in our models.

Of particular note is the investigated differences in self-regulation between high and low performing students in a rather selective sub-sample. Specifically, all participants in the present study acquired the formal qualification for a university entrance. This limits the lower end of the grades we investigated. Still, we found sizable differences in self-regulation within this sample. A potential avenue for further investigations is to extend our approach to a broader sample, including students from lower track(s) and students who failed to meet the requirements for university entry. Such an extension of the study sample seems to be particularly fruitful to relate investigations of self-regulation in educational contexts to similar investigations in other areas (e.g., personality or clinical psychology).

The modelling approach we chose entailed some limitations as well. We focused on explainable and parsimonious models to predict grades and laboratory learning outcomes and avoided potential gains in prediction accuracy that more complex and opaque (non-linear) models might be able to achieve. For instance, preliminary tests indicated that a less strict feature selection might have led to more accurate predictions of laboratory learning outcomes. However, from a psychological perspective we decided against such approaches as the interpretation of the models would have become increasingly difficult. Moreover, we did not investigate potential interactions among predictors. Research has shown that some of the variables we included can be moderators for other constructs (Bidjerano & Daj, 2007; Chamorro-Premuzic, & Furnham, 2003; Schunk & Greene, 2018) and the effectiveness of some constructs can best be assessed through their interaction with other variables (e.g., conscientiousness and interest, Trautwein et al., 2015). However, to maintain explainable models and due to the computational cost of exploring interactions between 39 predictors, we focused on the ‘main effects’ of predictors and their additive predictive value in our analyses.

All in all, to our knowledge, our study has provided the first integrative, empirical investigation of multiple areas of research on self-regulation in educational settings. Specifically, by using machine learning approaches we were able to show that predicting academic achievement in schools and laboratory learning outcomes relies on distinct patterns of self-regulatory constructs. While the most important predictors – motivation and working-memory capacity – were important for both outcomes, further central predictors varied significantly depending on the outcome. We further found that optimal predictions of learning consistently relied on predictors from all four areas of research on self-regulation in education (i.e., learning activities, driving forces, limited resources, and personal dispositions), showcasing their complementary value for learning. Therefore, the present study builds groundwork for further integrative studies on self-regulation by showing that learning in school and in the laboratory have different requirements in self-regulation but can be connected at multiple levels. Future research building upon these findings is required to disentangle the manifold constructs that are investigated under the umbrella term of self-regulation in education.

References

Abler, B., & Kessler, H. (2009). Emotion regulation questionnaire–Eine deutschsprachige Fassung des ERQ von Gross und John. *Diagnostica*, *55*(3), 144–152.

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*(1), 20–29.

Azevedo, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? The role of self-regulated learning. *Educational Psychologist*, *40*(4), 199–209.

Azevedo, R., Harley, J., Trevors, G., Duffy, M., Feyzi-Behnagh, R., Bouchet, F., & Landis, R. (2013). Using trace data to examine the complex roles of cognitive, metacognitive, and emotional self-regulatory processes during learning with multi-agent systems. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 427–449). Amsterdam, The Netherlands: Springer.

Azevedo, R., Taub, M., & Mudrick, N. (2018). Understanding and reasoning about real-timecognitive, affective, and metacognitive processes to foster self-regulation with advancedlearning technologies. In D. H. Schunk & J. A. Greene (Eds.).Handbook of self-regulation oflearning and performance(2nd ed., pp. 254–270). New York, NY: Routledge.

Baumeister, R. F., Heatherton, T. F., & Tice, D. M. (1994). *Losing control: How and why people fail at self-regulation.* Academic Press.

Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science, 16*(6), 351–355. https://doi.org/10.1111/j.1467-8721.2007.00534.x

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, *57*, 289–300.

Bertrams, A., & Dickhäuser, O. (2009). Messung dispositioneller selbstkontroll-kapazität: eine deutsche adaptation der kurzform der self-control scale (SCS-KD). *Diagnostica*, *55*(1), 2-10. https://doi.org/10.1026/0012-1924.55.1.2

Bidjerano, T., & Dai, D. Y. (2007). The relationship between the Big-Five model of personality and self-regulated learning strategies*. Learning and Individual Differences*, *17*, 69–81.

Boekaerts, M. (1996). Self-regulated learning at the junction of cognition and motivation. *European Psychologist*, *1*(2), 100–112. https://doi.org/10.1027/1016-9040.1.2.100

Boekaerts, M., & Pekrun, P. R. (2016). Emotion and emotion regulation in academic settings. In L. Corno & E. M. Anderman (Eds.), *Handbook of educational psychology* (pp.76–90). New York, NY: Routledge.

Boerner, S., Seeber, G., Keller, H., & Beinborn, P. (2005). Lernstrategien und lernerfolg im studium. *Zeitschrift Für Entwicklungspsychologie Und Pädagogische Psychologie*, *37*(1), 17–26.

Burnette, J. L., O'Boyle, E. H., VanEpps, E. M., Pollack, J. M., & Finkel, E. J. (2013). Mind-sets matter: A meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin, 139*(3), 655–701. https://doi.org/10.1037/a0029531

Chamorro-Premuzic, T., & Furnham, A. (2008). Personality, intelligence and approaches to learning as predictors of academic performance. *Personality and individual differences*, *44*(7), 1596-1603.

Costa, P. T., & McCrae, R. R. (1998). Six approaches to the explication of facet-level traits: examples from conscientiousness. *European Journal of Personality*, *12*(2), 117–134.

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, *26*, 197–223. https://doi.org/10.1007/s10648-013-9246-y

Credé, M., Tynan, M. C., & Harms, P. D. (2017). Much ado about grit: a meta-analytic synthesis of the grit literature. *Journal of Personality and social Psychology*, *113*(3), 492.

Daneman, M. , & Carpenter, P.A . (1980). Individual differences in working memory and reading. Journal of Verbal Learning an d Verbal Behavior,19, 450±466

Dang, J., King, K. M., & Inzlicht, M. (2020). Why are self-report and behavioral measures weakly correlated? *Trends in Cognitive Sciences*, *24*(4), 267–269.

de Bruin, A. B. H., & van Merriënboer, J. J. G. (2017). Bridging cognitive load and self-regulated learning research: A complementary approach to contemporary issues in educational research. *Learning and Instruction*, *51*, 1–9.

Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review*, *28*(3), 425–474. https://doi.org/10.1007/s10648-015-9320-8

Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, *3*(3), 231–264.

Dignath, C., Buettner, G., & Langfeldt, H.-P. (2008). How can primary school students learn self-regulated learning strategies most effectively?: A meta-analysis on self-regulation training programmes. *Educational Research Review*, *3*(2), 101–129.

Duckworth, A. (2016). *Grit: The power of passion and perseverance.* Scribner/Simon & Schuster.

Duckworth, A. L., & Kern, M. L. (2011). A meta-analysis of the convergent validity of self-control measures. *Journal of research in personality*, *45*(3), 259-268.

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the Short Grit Scale (GRIT–S). *Journal of personality assessment*, *91*(2), 166-174.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*(1), 109–132.

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist*, *46*(1), 6–25. https://doi.org/10.1080/00461520.2011.538645

Eilam, B., Zeidner, M., & Aharon, I. (2009). Student conscientiousness, self-regulated learning, and science achievement: An explorative field study. *Psychology in the Schools*, *46*, 420–432.

Eisenberg, I. W., Bissett, P. G., Enkavi, A. Z., Li, J., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Uncovering the structure of self-regulation through data-driven ontology discovery. *Nature Communications*, *10*(1), 1–13.

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, *113*(28), 7900–7905.

Ekstrom, R., French, J., Harman, H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests.* Princeton: Educational Testing Service.

Ent, M. R., Baumeister, R. F., & Tice, D. M. (2015). Trait self-control and the avoidance of temptation. *Personality and Individual Differences, 74,* 12–15. https://doi.org/10.1016/j.paid.2014.09.031

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186–204.

Friso-van den Bos, I., van der Ven, S. H. G., Kroesbergen, E. H., & van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review*, *10*, 29–44. https://doi.org/https://doi.org/10.1016/j.edurev.2013.05.003

Gerjets, P., Scheiter, K., & Catrambone, R. (2006). Can learning from molar and modular worked examples be enhanced by providing instructional explanations and prompting self-explanations?. *Learning and Instruction*, *16*(2), 104-121.

Greene, J. A., & Azevedo, R. (2009). A macro-level analysis of SRL processes and their relations to the acquisition of a sophisticated mental model of a complex system. *Contemporary Educational Psychology*, *34*(1), 18–29.

Greene, J. A., Bolick, C. M., Jackson, W. P., Caprino, A. M., Oswald, C., & McVea, M. (2015). Domain-specificity of self-regulated learning processing in science and history. *Contemporary Educational Psychology*, *42*, 111–128.

Gross, J. J. (2013). Emotion regulation: Taking stock and moving forward. *Emotion*, *13*, 359–365. https://doi.org/10.1037/a0032135

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*(2), 348-362.

Gustavson, D. E., Miyake, A., Hewitt, J. K., & Friedman, N. P. (2015). Understanding the cognitive and genetic underpinnings of procrastination: Evidence for shared genetic influences with goal management and executive function abilities. *Journal of Experimental Psychology: General, 144*(6), 1063–1079. https://doi.org/10.1037/xge0000110

Gustavson, D. E., Stallings, M. C., Corley, R. P., Miyake, A., Hewitt, J. K., & Friedman, N. P. (2017). Executive functions and substance use: Relations in late adolescence and early adulthood. *Journal of Abnormal Psychology, 126*(2), 257–270. https://doi.org/10.1037/abn0000250

Haatveit, B. C., Sundet, K., Hugdahl, K., Ueland, T., Melle, I., & Andreassen, O. A. (2010). The validity of d prime as a working memory index: results from the “Bergen n-back task”. *Journal of clinical and experimental neuropsychology*, *32*(8), 871-880.

Harley, J. M., Pekrun, R., Taxer, J. L., & Gross, J. J. (2019). Emotion regulation in achievement situations: An integrated model. *Educational Psychologist*, *54*(2), 106–126.

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

Hill, P. L., Nickel, L. B., & Roberts, B. W. (2014). Are you in a healthy relationship? Linking conscientiousness to health via implementing and immunizing behaviors. *Journal of Personality*, *82*(6), 485–492.

Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction, 17*(6), 722–738. https://doi.org/10.1016/j.learninstruc.2007.09.013

Hofmann, W., Baumeister, R. F., Förster, G., & Vohs, K. D. (2012). Everyday temptations: An experience sampling study of desire, conflict, and self-control. *Journal of Personality and Social Psychology, 102*(6), 1318–1335. https://doi.org/10.1037/a0026545

Inzlicht, M., Werner, K. M., Briskin, J. L., & Roberts, B. W. (2021). Integrating models of self-regulation. *Annual Review of Psychology*, *72*, 319–345.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), e124.

Jachimowicz, J. M., Wihler, A., Bailey, E. R., & Galinsky, A. D. (2018). Why grit requires perseverance and passion to positively predict performance. *Proceedings of the National Academy of Sciences*, *115*(40), 9980-9985.

Jacobson, M. J., & Spiro, R. J. (1995). Hypertext learning envir.onments, cognitive flexibility, and the transfer of complex knowledge: An empirical investigation. *Journal of Educational Computing Research, 12*, 301-333. https://doi.org/10.2190/4T1B-HBP0-3F7E-J4PN

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.

Kornmann, J., Kammerer, Y., Zettler, I., Trautwein, U., & Gerjets, P. (2016). Hypermedia exploration stimulates multiperspective reasoning in elementary school children with high working memory capacity: A tablet computer study. *Learning and Individual Differences*, *51*, 273-283.

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.

Kunter, M., Schümer, G., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., . . . Weiß, M. (2002). *PISA 2000: Dokumentation der Erhebungsinstrumente* [PISA 2000: Documentation of Scales]. Berlin, Germany: Heenemann GmbH.

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

Levens, S. M., & Gotlib, I. H. (2012). The effects of optimism and pessimism on updating emotional information in working memory. *Cognition & emotion*, *26*(2), 341-350.

Loderer, K., Pekrun, R., & Lester, J. C. (2020). Beyond cold technology: A systematic review and meta-analysis on emotions in technology-based learning environments. *Learning and Instruction*, *70*, 101162. https://doi.org/https://doi.org/10.1016/j.learninstruc.2018.08.002

Logan, G. D. (1994). On the ability to inhibit thought and action: A user's guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory Processes in Attention, Memory and Language* (pp. 189–239). San Diego: Academic Press.

Malanchini, M., Engelhardt, L. E., Grotzinger, A. D., Harden, K. P., & Tucker-Drob, E. M. (2019). "Same but different": Associations between multiple aspects of self-regulation, cognition, and academic abilities. *Journal of Personality and Social Psychology, 117*(6), 1164–1188. https://doi.org/10.1037/pspp0000224

Marsh, H. W. (1990). The structure of academic self-concept: The Marsh/Shavelson model. *Journal of Educational Psychology, 82*(4), 623–636. https://doi.org/10.1037/0022-0663.82.4.623

Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, *81*(1), 59-77.

Matzke, D., Dolan, C. V., Logan, G. D., Brown, S. D., & Wagenmakers, E. J. (2013). Bayesian parametric estimation of stop-signal reaction time distributions. *Journal of Experimental Psychology: General*, *142*(4), 1047-1073.

Mayer, R. E. (2014). Multimedia instruction. InHandbook of research on educational communications and technology (pp. 385-399). New York, NY: Springer.

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., … Ross, S. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, *108*(7), 2693–2698.

Morell, M., Yang, J. S., Gladstone, J. R., Turci Faust, L., Ponnock, A. R., Lim, H. J., & Wigfield, A. (2020). Grit: The long and short of it. *Journal of Educational Psychology.* Advance online publication. https://doi.org/10.1037/edu0000594

Muenks, K., Wigfield, A., Yang, J. S., & O'Neal, C. R. (2017). How true is grit? Assessing its relations to high school and college students' personality characteristics, self-regulation, engagement, and achievement. *Journal of Educational Psychology*, *109*(5), 599–620. https://doi.org/10.1037/edu0000153

Namkung, J. M., Peng, P., & Lin, X. (2019). The relation between mathematics anxiety and mathematics performance among school-aged students: a meta-analysis. *Review of Educational Research*, *89*(3), 459–496.

Narciss, S., Proske, A., & Koerndle, H. (2007). Promoting self-regulated learning in web-based learning environments. *Computers in Human Behavior*, *23*(3), 1126–1144. https://doi.org/https://doi.org/10.1016/j.chb.2006.10.006

National Research Council. (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. Washington, DC: The National Academies Press. https://doi.org/10.17226/13398

Neurobehavioral Systems, Inc. [Presentation 18.0]. (2018). Retrieved from https://www.neurobs.com/

Nigg, J. T. (2017). Annual Research Review: On the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychopathology. *Journal of child psychology and psychiatry*, *58*(4), 361-383. https://doi.org/10.1111/jcpp.12675

OECD. (2013). *Trends Shaping Education 2013*. Paris. France: OECD. https://doi.org/10.1787/trends_edu-2013-en

Ostendorf, F., & Angleitner, A. (2004). *NEO-Persönlichkeitsinventar nach Costa und McCrae, revidierte Form (NEO-PI-R). [NEO Personality Inventory based on Costa and McCrae, revised version (NEO-PI-R)]*. Göttingen, Germany: Hogrefe.

Paas, F., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. *Educational Psychology Review*, *24*(1), 27–45. https://doi.org/10.1007/s10648-011-9179-2

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology, 8,* Article 422. https://doi.org/10.3389/fpsyg.2017.00422

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

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). San Diego, CA: Academic Press. https://doi.org/10.1016/B978-012109890-2/50043-3

Pintrich, P. R. (2003). A Motivational Science Perspective on the Role of Student Motivation in Learning and Teaching Contexts. *Journal of Educational Psychology*, *95*(4), 667–686. https://doi.org/10.1037/0022-0663.95.4.667

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.

Ponnock, A., Muenks, K., Morell, M., Seung Yang, J., Gladstone, J. R., & Wigfield, A. (2020). Grit and conscientiousness: Another jangle fallacy. *Journal of Research in Personality*, *89*, 104021. https://doi.org/https://doi.org/10.1016/j.jrp.2020.104021

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*(2), 322–338. https://doi.org/10.1037/a0014996

Puustinen, M., & Pulkkinen, L. (2001). Models of self-regulated learning: A review. *Scandinavian Journal of Educational Research*, *45*(3), 269–286.

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

Rieger, S., Göllner, R., Spengler, M., Trautwein, U., Nagengast, B., &Roberts, B. W. (2017). Social cognitive constructs are just as stable asthe big five between grades 5 and 8.AERA Open, 3,1–9.http://dx.doi.org/10.1177/2332858417717691

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, *2*(4), 313–345.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin, 132*(1), 1–25. https://doi.org/10.1037/0033-2909.132.1.1

Robson, D. A., Allen, M. S., & Howard, S. J. (2020). Self-regulation in childhood as a predictor of future outcomes: A meta-analytic review. *Psychological Bulletin, 146*(4), 324–354. https://doi.org/10.1037/bul0000227

Saeys, Y., Inza, I., & Larranaga, P. (2007). A review of feature selection techniques in bioinformatics. *bioinformatics*, *23*(19), 2507-2517. https://doi.org/10.1093/bioinformatics/btm344

Schneider, W., Schlagmüller, M., & Ennemoser, M. (2017). *LGVT 5-12+: Lesegeschwindigkeits- und -verständnistest für die Klassen 5-12: Manual [Reading speed and comprehesion test for grades 5-12: Manual]*. Göttingen, Germany: Hogrefe.

Schunk, D. H., & Greene, J. A. (2018). Handbook of Self-Regulation of Learning and Performance. In *Handbook of Self-Regulation of Learning and Performance* (2nd Edition). New York: Routledge. https://doi.org/10.4324/9781315697048

Shanahan, M. J., Hill, P. L., Roberts, B. W., Eccles, J., & Friedman, H. S. (2014). Conscientiousness, health, and aging: The Life Course of Personality Model. *Developmental Psychology, 50*(5), 1407–1425. https://doi.org/10.1037/a0031130

Sitzmann, T., & Ely, K. (2011). A meta-analysis of self-regulated learning in work-related training and educational attainment: What we know and where we need to go. *Psychological Bulletin, 137*(3), 421–442. https://doi.org/10.1037/a0022777

Psychology Software Tools, Inc. [E-Prime 3.0]. (2018). Retrieved from https://support.pstnet.com/.

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of personality*, *72*(2), 271-324.

Taub, M., Sawyer, R., Lester, J., & Azevedo, R. (2020). The impact of contextualized emotions on self-regulated learning and scientific reasoning during learning with a game-based learning environment. *International Journal of Artificial Intelligence in Education*, *30*(1), 97-120.

The World Bank Group. (2011). 2011). *Learning for all: Investing in people's knowledge and skills to promote development: Education sector strategy 2020*. Washington, DC: The World Bank. https://doi.org/10.1086/675413

Trautwein, U. (2007). The homework–achievement relation reconsidered: Differentiating homework time, homework frequency, and homework effort. *Learning and Instruction*, *17*(3), 372-388.

Trautwein, U., Lüdtke, O., Nagy, N., Lenski, A., Niggli, A., & Schnyder, I. (2015). Using individual interest and conscientiousness to predict academic effort: Additive, synergistic, or compensatory effects? *Journal of Personality and Social Psychology, 109*(1), 142–162. https://doi.org/10.1037/pspp0000034

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent?. *Journal of memory and language*, *28*(2), 127-154.

Tze, V. M. C., Daniels, L. M., & Klassen, R. M. (2016). Evaluating the relationship between boredom and academic outcomes: A meta-analysis. *Educational Psychology Review*, *28*(1), 119–144.

van Genugten, L., Dusseldorp, E., Massey, E. K., & van Empelen, P. (2017). Effective self-regulation change techniques to promote mental wellbeing among adolescents: a meta-analysis. *Health Psychology Review*, *11*(1), 53-71. https://doi.org/10.1080/17437199.2016.1252934

van Merriënboer, J. J. G., & Kirschner, P. A. (2007). *Ten steps to complex learning: A systematic approach to four-component instructional design.* Lawrence Erlbaum Associates Publishers.

Weiß, R. H. (2006). CFT 20-R. *Grundintelligenztest Skala 2 – Revision [Cultural fair intelligence test 2 - Revision]*. Göttingen, Germany: Hogrefe.

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: issues and guidance for practice. *Statistics in medicine*, *30*(4), 377-399.

Wild, K.-P. and Schiefele, A. (1994). Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens [Learning strategies at university: Findings on the factorial structure and reliability of a new questionnaire]. *Zeitschrift fuer Differentielle und Diagnostische Psychologie*,*15*, 185/120.

Winne, P. H., & Hadwin, A. (2008). The weave of motivation and self-regulated learning. In D. Schunk & B. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 297–314). Mahwah, NJ: Erlbaum.

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

Winne, P. H., & Perry, N. E. (2000). *Measuring self-regulated learning.* In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (p. 531–566). Academic Press. https://doi.org/10.1016/B978-012109890-2/50045-7

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*, *12*(6), 1100-1122.

Zheng, L. (2016). The effectiveness of self-regulated learning scaffolds on academic performance in computer-based learning environments: A meta-analysis. *Asia Pacific Education Review*, *17*(2), 187–202.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In *Handbook of self-regulation.* (pp. 13–39). San Diego, CA, US: Academic Press. https://doi.org/10.1016/B978-012109890-2/50031-7

Appendix 2A

Table 2.5.

Overview of scales and subscales of self-report measures used in the present study

Questionnaire	Subscales	Items	α	References
Neo Personality Inventory	Agreeableness	48	.90	Costa & McCrae, 2008; Ostendorf & Angleitner, 2004
	Conscientiousness	48	.92	
	Extraversion	48	.90	
	Neuroticism	48	.93	
	Openness	48	.86	
Brief Self-Control Scale	-	13	.83	Bertrams & Dickhäuser, 2009; Tangney et al., 2004
Inventory for the Measurement of Learning Strategies in Academic Studies	Organization	9	.67	Boerner et al., 2005; Wild & Schiefele, 1994
	Relationships	8	.84	
	Critical Evaluation	8	.88	
	Rehearsal	8	.81	
	Effort	8	.81	
	Attention	6	.93	
	Time Management	4	.87	
	Learning Environment	6	.83	
	Learning with fellow students	4	.80	
	Literature	4	.80	
	Metacognitive Strategies	20	.86	
Emotion Regulation Questionnaire	Cognitive Reappraisal	6	.81	Abler & Kessler, 2009
	Expressive Suppression	4	.75	
GRIT	Consistency of Interest	7	.79	Morell et al., 2020
	Perseverance of Effort	7	.81	
Cross-Curricular Competencies	Control Strategies	5	.58	Kunter et al., 2002
	Elaboration Strategies	4	.82	
	Effort and Persistence in	4	.76	
	Learning	4	.77	
	Self-Efficacy	3	.91	
	Instrumental Motivation			
Subject Specific Motivation (Self-Concept, Interest, Value)	Self-Concept Math	4	.93	
	Self-Concept History	4	.89	
	Self-Concept Art	4	.89	
	Self-Concept Physics	4	.91	
	Self-Concept Biology	4	.88	
	Interest Math	2	.94	
	Interest History	2	.93	
	Interest Art	2	.93	
	Interest Physics	2	.92	
	Interest Biology	2	.93	
	Value Math	4	.88	
	Value History	4	.83	
	Value Art	4	.91	
	Value Physics	4	.91	
	Value Biology	4	.89	

Appendix 2B

Table 2.6.

Overview of cognitive task data used in the present study

Task/Paradigm	Measure	References
CFT	Sum of correct responses	(Weiß, 2006)
N-Back Task	d-prime score (d’)	(Levens & Gotlib, 2012)
Operation Span Task	Maximum span reached in adaptive O-Span task	(Turner & Engle, 1989)
PFT	Sum of correct responses	(Ekstrom, Dermen, & Harman, 1976)
Reading Comprehension Test	Comprehension Accuracy	(Schneider, Schlagmüller, & Ennemoser, 2017)
Reading Span Task	Maximum span reached in adaptive R-Span task	(Turner & Engle, 1989)
Stop Signal Task	Stop Signal Reaction Time	(Bissett & Logan, 2011)
Stroop Task	Reaction time difference for Color Word Interference Print Color Interreference	(MacLeod, 2005)
Switching Task	Reaction time difference between switch and non-switch trials Difference in accuracy between switch and non-switch trials	(Sudevan & Taylor, 1987)
Wisconsin Card Sorting Task	Count of omission errors Total count of Errors	(Grant & Berg, 1948)

Appendix 2C

Table 2.7.

Mean values and standard deviations for the top ten features of grade models by performance group

Predictor	Grades			
	Top 30%		Bottom 30%	
	M	SD	M	SD
Mean Expectancy Value	12.49	2.11	8.93	2.24
Effort (LIST)	2.78	0.74	2.26	0.63
D Prime (3-back)	2.03	0.71	1.55	0.88
D Prime (2-back)	2.63	1.06	1.94	1.16
Control Strategies	2.40	0.51	2.19	0.41
Effort and Persistence in Learning	2.34	0.60	2.00	0.53
Reading Span	5.99	1.08	5.55	0.95
Mean Interest	2.78	0.46	2.29	0.44
Conscientiousness	128.12	23.00	116.16	22.41
Self-Control (BSCS)	2.36	0.67	2.04	0.57

Table 2.8.

Mean values and standard deviations for the top ten features of laboratory task performance models by performance group

Predictor	Laboratory Learning Task Performance			
	Top 30%		Bottom 30%	
	M	SD	M	SD
Rehearsal strategies (LIST)	2.10	0.75	2.67	0.69
D Prime (3-back)	2.03	0.84	1.47	0.72
Mean Expectancy Value	11.47	2.29	9.86	2.70
Openness	133.57	16.97	121.17	18.08
Switching Accuracy Difference	-0.03	0.03	-0.04	0.04
Reading Span	6.16	0.76	5.66	0.95
D Prime (2-back)	2.49	1.50	1.71	1.16
Time Management (LIST)	1.55	1.05	1.99	1.03
Memorization Strategies (CCC)	1.22	0.73	1.78	0.71
Mean Interest	2.68	0.46	2.43	0.45

3 Study II

To click or to touch? Learning environments based on tablet versus personal computers differ in the self-regulation requirements imposed onto learners

Franz Wortha, Philipp Mock, Birgit Brucker, Maike Tibus, Olga Özbek & Peter Gerjets

Note. Unpublished Manuscript. This article might not exactly replicate the final version published in the journal.

Estimated contributions

Scientific ideas by the candidate (%)	Data generation by the candidate (%)	Analysis and Interpretation by the candidate (%)	Paper writing done by the candidate (%)
80	30	90	90

Abstract

Touch-devices have become the primary means of interaction with information technology. This development led to a significant increase in the use of touch-devices (i.e., tablets) in educational settings. Accordingly, research has started to investigate how different affordances of tablet devices affect learning and related processes. One line of research in this context identified strong links between tablet use and motivational facets of learning, particularly perceived effort requirements. Further research has shown that specific properties of interactions with tablets, including the gestures used to control them and the processing of information near the hands might have profound effects on cognitive processes and learning in turn. Both lines of research have shown potential beneficial and detrimental effects of tablet use for learning, but the state of research remains fragmented. A promising approach to overcome this issue lies in the use of self-regulation as a conceptual framework. An extensive amount of research on self-regulation in educational contexts, including research on self-regulated learning, motivational aspects of learning, cognitive resources, and personal predispositions, has shown that successful learning requires the learner to adequately adapt their thoughts, behaviors, actions, and feelings to the learning task they are pursuing. Despite the widespread use of touch-devices in educational settings, little research has directly investigated the importance of self-regulation for learning on tablets in contrast to PCs. In this study, we addressed this issue by investigating a broad set of self-regulatory constructs and their predictive value in an art learning task either performed on a tablet or a PC. Results indicated that differences in self-regulation between high and low performer were more pronounced for students using tablets compared to learners using PCs. Specifically, machine learning models showed that the use of tablets was associated with higher self-regulatory demands across multiple facets of self-regulation during learning. Implications of these findings for the design and development of learning tasks on tablets were discussed.

Introduction

Over the last decade, touch-based interaction has become the most common form of interaction with information technology (Oviatt & Cohen, 2015). With this development the effective use of touch devices (particularly tablets) in educational contexts has become an important opportunity but also challenge for practitioners and researchers alike. A recent surge of research has started to investigate potential beneficial and detrimental effects that the use of touch-based interaction can have on learning (Haßler, Major, & Hennessy, 2016; Mulet, Van De Leemput, & Amadieu, 2019; Turvey & Pachler, 2018). Overall, these examinations showed that the use of tablets commonly has positive effects on motivation and learning, but that further research is needed to comprehensively understand the impact of touch-interactions in educational settings. Further, research focusing on how cognitive functions (e.g., attention, cognitive control) are affected by specific affordances inherent to touch-interactions (e.g., hand proximity, use of gestures) demonstrated great potentials of touch interactions to enhance learning and related processes (Brucker, Brömme, Ehrmann, Edelmann, & Gerjets, 2021; Weidler & Abrams, 2014). Yet, these two research traditions remain largely independent from one another. Moreover, so far, these facets of the use of touch-devices in educational contexts have not systematically considered potential connections to further central traditions of educational research, such as studies on learning activities (e.g., Dent & Koenka, 2016) or personal dispositions (e.g., Poropat, 2009). This integration on the other hand is crucial to use the full potential of touch-based technology (e.g., tablets) for education.

3.1.1 Learning with tablets: Ease of use and amount of invested mental effort

Touch-devices such as tablets a are becoming increasingly popular (Oviatt & Cohen, 2015) in schools and other educational contexts due their media-specific characteristics that render them particularly promising for different learning scenarios across content domains and learning tasks (Haßler et al., 2016). Their advantages when compared to other learning mediums (including PCs) involve their ease and flexibility of use (e.g., availability, portability, intuitive touch interaction, large number of simple applications for single and shared use, low threshold for the reception and production of digital contents). A recent review has shown that tablets' capability to foster learning across domains is supported empirically (Mulet et al., 2019). However, these investigations also indicated that the beneficial effect of tablets may vary depending on characteristics of the learning environment and task. For instance, Mulet et al. (2019) outlined that the regulation of learning processes in terms of teacher guidance is an important aspect to ensure positive learning experiences and outcomes for students. Moreover, the review also confirmed the assumption that participants generally perceived tablets as easy to use, which in turn leads to lower levels of effort expectancy, which can, however, also be seen as a rather critical aspect. For instance, Schwab et al. (Schwab, Hennighausen, Adler, & Carolus, 2018) emphasized the importance of perceived effort demands for learning with touch devices in a recent review of research on Salomon's AIME model (amount of invested mental effort). (Salomon, 1984). According to this model, students' perceived demand characteristics and self-efficacy are the crucial factors for the amount of effort they invest during learning when no specific guidelines for effort investment are provided. Therefore, if students perceive a medium as imposing rather low effort demands, they will also invest less mental effort, which potentially negatively affects learning outcomes. This could imply that successfully learning with an "easy" tablet might even require more self-regulatory processes to overcome the perceived "easiness" of touch devices when compared to traditional PCs what might be perceived as "harder" to learn with. Without these self-regulatory processes, a lack of invested mental effort might even counteract the potential positive characteristics of tablet devices.

Touch-devices such as tablets a are becoming increasingly popular (Oviatt & Cohen, 2015) in schools and other educational contexts due their media-specific characteristics that render them particularly promising for different learning scenarios across content domains and learning tasks (Haßler et al., 2016). Their advantages when compared to other learning mediums (including PCs) involve their ease and flexibility of use (e.g., availability, portability, intuitive touch interaction, large number of simple applications for single and shared use, low threshold for the reception and production of digital contents). A recent review has shown that tablets' capability to foster learning across domains is supported empirically (Mulet et al., 2019). However, these investigations also indicated that the beneficial effect of tablets may vary depending on characteristics of the learning environment and task. For instance, Mulet et al. (2019) outlined that the regulation of learning processes in terms of teacher guidance is an important aspect to ensure positive learning experiences and outcomes

for students. Moreover, the review also confirmed the assumption that participants generally perceived tablets as easy to use, which in turn leads to lower levels of effort expectancy, which can, however, also be seen as a rather critical aspect. For instance, Schwab et al. (Schwab, Hennighausen, Adler, & Carolus, 2018) emphasized the importance of perceived effort demands for learning with touch devices in a recent review of research on Salomon's AIME model (amount of invested mental effort). (Salomon, 1984). According to this model, students' perceived demand characteristics and self-efficacy are the crucial factors for the amount of effort they invest during learning when no specific guidelines for effort investment are provided. Therefore, if students perceive a medium as imposing rather low effort demands, they will also invest less mental effort, which potentially negatively affects learning outcomes. This could imply that successfully learning with an "easy" tablet might even require more self-regulatory processes to overcome the perceived "easiness" of touch devices when compared to traditional PCs what might be perceived as "harder" to learn with. Without these self-regulatory processes, a lack of invested mental effort might even counteract the potential positive characteristics of tablet devices.

3.1.2 Touch interactions, gestures and cognitive functions

Investigations based on cognitive load theory (Paas & Sweller, 2012) and on embodied cognition (Foglia & Wilson, 2013) showed that (touch) gestures – the primary form of interaction with tablets – can be beneficial for learning. According to these theoretical approaches gestures have the potential to benefit learning because they (and movements in general) may help guiding attention, thereby freeing up working memory resources which in turn can be deployed for other task-related processes (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001; Sepp, Agostinho, Tindall-Ford, & Paas, 2020). In line with this reasoning, research indicated that different types of gestures can indeed foster learning. For instance, multiple studies have shown that tracing gestures have a positive effect in different learning tasks, (Agostinho et al., 2015; Ginns, Hu, Byrne, & Bobis, 2016; Hu, Ginns, & Bobis, 2015). As an example, Agostinho and colleagues (Agostinho et al., 2015) found that tracing elements of worked examples in math learning was beneficial for learning in primary school students. Beyond the reduction of cognitive load gestures can further affect learning processes through their impact on attention, perception or memory (Cook, 2018). These benefits have been shown in multiple learning related context, such as problem-solving tasks (Chu & Kita, 2011) or the understanding of dynamic systems (Kang & Tversky, 2016).

Another line of research indicates that interaction with touch devices can alter learning and related processes through the mere presence of the hands near the materials. Specifically, potential benefits of touch interactions regarding the deployment of attentional resources and cognitive control can be derived from research investigating the effects of hand proximity on attentional and cognitive processing. This research indicates that the peripersonal space around the hands plays a unique role in information processing. Generally, findings in this field of research show that processing of visuospatial information is enhanced and more thoroughly when stimuli are located close to one's hands (Abrams, Weidler, & Suh, 2015; Tseng & Bridgeman, 2011). For instance, studies showed that hand proximity is associated with deeper levels of processing (e.g., change detection performance) for pictorial information (Tseng & Bridgeman, 2011). The processing of textual information, however, does not seem to benefit from hand proximity or might even be impaired close to the hands (Davoli, Du, Montana, Garverick, & Abrams, 2010). A recent study showed that these findings transfer to learning contexts, as learning of visuospatial materials on a multi-touch table benefitted from hand proximity, whereas no such effects for verbal learning occurred (Brucker et al., 2021). In summary, the interaction with touch-devices can affect a multitude of learning-related processes through its innate proximity of the hands to the material and use of gestures. Particularly for visuospatial materials tablets might provide novel digital opportunities to foster learning due to their potential for guiding attention, enhancing visual information processing, and supporting cognitive control processes (e.g., by offloading cognitive load). From a broader perspective, they can support cognitive resources required in complex, self-regulated learning tasks. For instance, by freeing up working memory more resources are available to engage in sophisticated learning activities and their regulation.

Taken together, the two lines of research outlined above indicate that the use of touch-devices (i.e., tablets) in educational settings has great potential to foster learning from multiple angles. Primarily, these investigations have focused on motivational forces that can drive the learning activity (e.g., perceived mental effort; Mulet et al., 2019) or underlying cognitive resources required for learning (e.g., working memory; Sepp et al., 2020). However, how these different perspectives interact has not been investigated. Further, research from different backgrounds has shown that learning and academic achievement are not only driven by driving motivational forces or underlying cognitive resources, but are also dependent on regulation of learning activities (e.g., cognitive and metacognitive learning strategies) and students' personal dispositions (e.g., conscientiousness; Richardson, Abraham, & Bond, 2012). In the following sections we will outline the landscape of research on these constructs (i.e., learning activates, driving forces, personal dispositions, and limited resources) carried out under the umbrella term of self-regulation. Then we will demonstrate the importance of investigating the effect of touch interactions for learning in this broad context of self-regulation in education.

3.1.3 Self-regulation in education

The ability to self-regulate ones learning and related processes is a central skill necessary to meet the requirements and challenges that contemporary and future educational settings impose, such as for example the widespread use of (complex) information technology in educational settings (OECD, 2013; Pellegrino & Hilton, 2012; The World Bank Group, 2011).

Self-regulation (SR) in educational settings can be broadly defined as processes that involve planning, monitoring, and controlling thoughts, behaviors, emotions, and motivation in pursuit of goals (Schunk & Greene, 2018). Under this umbrella term research from multiple research traditions, including cognitive, educational or social psychology, education research, and cognitive science have investigated different aspects of SR and their relation to learning, learning outcomes, and academic achievement. The focus of these lines of research varies greatly and spans from investigations of (meta-)cognitive processes when learning with advanced learning technologies, to the importance of cognitive resources for learning in laboratory studies and school, and investigations of the impact of rather stable personal dispositions that affect important life outcomes. In the following sections, we will briefly introduce four central themes of research related to SR, specifically self-regulated learning (SRL), emotional and motivational influences, limitations of cognitive resources, and personal dispositions of the learners. We will use these perspectives to examine if and how self-regulatory requirements in an art learning task are different depending on the modality (i.e., learning with a PC vs. learning with a tablet).

(1) Learning Activities. One of the most prominent lines of research that directly investigates self-regulatory processes during learning is research on SRL. This field of inquiry typically investigates how specific cognitive, metacognitive and motivational processes unfold during learning activities and how they affect learning in laboratory settings, school, and university (e.g., Boekaerts & Corno, 2005; Winne & Hadwin, 1998, 2008). Multiple models emphasizing different aspects of SRL processes have been proposed and empirically tested (for a recent overview see Panadero, 2017). Two factors typically build the central pillars of these models – metacognitive monitoring and control (Nelson & Narens, 1994). More specifically, it is assumed that students constantly monitor properties of the learning process, such as the rate of progress towards the current goal and compare them to internal standards and goals (Butler & Winne, 1995; Carver & Scheier, 1998; Winne & Hadwin, 1998). Based on the outcome of this comparison appropriate learning strategies are deployed. Several metanalytic reviews support these assumptions as they have shown that cognitive and metacognitive learning processes, as well as interventions that aim to foster them are directly related to learning outcomes and academic achievement in schools, university, and vocational settings (Dent & Koenka, 2016; Dignath, Buettner, & Langfeldt, 2008; Dignath & Büttner, 2008; Sitzmann & Ely, 2011; L. Zheng, 2016).

(2) Driving Forces. Other research programs and models of SRL have extended the scope beyond cognitive and metacognitive processes by focusing on the forces driving SRL, such as emotional (e.g., Boekaerts & Pekrun, 2016; Efklides, 2011) or motivational (e.g., Pintrich, 2000, 2003; Zimmerman, 2000) influences. With regard to SRL theory, research on the motivational and emotional direction and energization of action provides a complementary approach to metacognitive monitoring and control (Pintrich, 2000). Outside of SRL research, several other research traditions on SR in educational, motivational, and social psychology have also investigated the regulation of motivation and emotion yet without strongly connecting it to SRL. This includes, for instance, situation-specific value beliefs and expectancies (Eccles & Wigfield, 2002), situational interests (Hidi & Renninger, 2006) or achievement emotions (Pekrun, 2006). Research on how emotions and motivations can be regulated, either in terms of influencing one's own situational (in contrast to dispositional) interests and motivational beliefs or in terms of using strategies for volitional and emotional control (Boekaerts & Niemivirta, 2000; Gross, 2013), has received much attention in many fields of research and needs to be more broadly connected to research on SRL.

(3) Limited Resources. Processes underlying SR require different kinds of limited processing resources, which are covered by research traditions different from the two lines of research outlined above. On the one hand, limitations of cognitive resources such as working memory capacity or executive functions (e.g., inhibiting the processing of irrelevant information) are important as they provide a basis for cognitive control and metacognition (Alloway & Alloway, 2010; Cowan, 2014; Miyake et al., 2000). Accordingly, strong relations between working memory, executive functions, and academic achievement have been found (Alloway & Alloway, 2010). Moreover, important instructional theories in educational psychology such as Cognitive Load Theory (e.g., Paas & Sweller, 2012) focus on the role of working-memory resources for learning and instruction. However, connections of these theories to SRL (e.g., de Bruin & van Merriënboer, 2017) or basic cognitive and neurocognitive research on working memory (Zheng, 2017) have only recently been attempted. On the other hand, attentional and volitional resources are needed to ensure goal achievement (e.g., for effort investment, regulation of emotions, and the willpower to resist impulses; Bauer & Baumeister, 2011; Gross, 2013; Mischel et al., 2011). Up to now, there are only a few recent attempts in this area to show conceptual and empirical associations between working memory and executive functions on the one hand and volitional resources on the other hand in terms of their role for self-regulated behavior (e.g., Hofmann, Friese, Schmeichel, & Baddeley, 2011; Hofmann, Schmeichel, & Baddeley, 2012; Ilkowska & Engle, 2010).

(4) Personal Disposition. Another line of research on SR in education focuses on the personal disposition a learner contributes to the situational context of an SRL scenario. In contrast to the first three core themes, research traditions studying dispositions address patterns of thoughts, feelings, and behaviors that are rather stable over time and situations and that provide more distal causes of SRL processes. Relevant dispositions for SR processes comprise personality traits (e.g., conscientiousness Costa & McCrae, 2008) but also other stable social cognitive constructs such as academic self-concepts and dispositional interests (Hidi & Renninger, 2006; Marsh, 1990), mindsets (Yeager & Dweck, 2012), or cognitive abilities (e.g., Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007). In contrast to driving forces (e.g., situational interest) and limited resources (e.g., willpower), personal dispositions are considered to be quite stable over situations. Individuals develop these dispositions for SR across the lifespan (Roberts, Walton, & Viechtbauer, 2006). In childhood, they are operationalized as temperament (Rothbart, 1989), that is, as an individual difference in SR and reactivity. In adolescence and adulthood, research on SR-relevant personality traits has focused on the Big Five trait of conscientiousness (Costa & McCrae, 1998), on grit (Duckworth, Peterson, Matthews, & Kelly, 2007), or on a combination of personality and stable motivation (Trautwein et al., 2015). Personal dispositions are relevant for SR in education as they are robust predictors of academic success (mediated by academic effort) that are fairly independent of cognitive ability (De Raad & Schouwenburg, 1996; Poropat, 2009) Yet, only few studies have investigated how personal dispositions are related to SRL. They have shown that conscientiousness was positively related to motivational aspects of learning, self-reported SRL strategy use, and achievement (Chamorro-Premuzic & Furnham, 2003; Eilam, Zeidner, & Aharon, 2009). Moreover, conscientiousness and openness were found to be associated with a more frequent use of metacognitive and elaborative learning strategies (Bidjerano & Dai, 2007). Neuroticism on the other hand is expected to have a negative relation to achievement as it is associated with emotional instability (Chamorro-Premuzic & Furnham, 2003).

However, although all of the aforementioned findings have shown great predictive value for learning outcomes (Dignath et al., 2008; Dignath & Büttner, 2008; Jansen, Van Leeuwen, Janssen, Jak, & Kester, 2019; Sitzmann & Ely, 2011), approaches to better integrate the constructs remain scarce (see Eisenberg et al., 2019 for an example). Considering a broader perspective on SR in education is particularly promising when investigating self-directed learning in complex digital learning environments that require different aspects of SR ranging from effort investment and resistance to distractions to the metacognitive planning and monitoring of learning activities, the maintenance of motivation and the engagement of cognitive resources when interacting with digital materials. In this paper, we will address a central aspect of learning with digital learning environments, that has been rarely considered in the context of SRL, the mode of interaction (i.e., learning with a PC vs. learning with a tablet). To this end, the following section will briefly outline how touch-interactions can potentially affect multiple facets of SR (i.e., learning activities, driving forces, personal dispositions, and limited resources).

3.1.4 SRL with tablets

It has been argued that touch-based learning environments provided, for instance, by tablets or smartphones might be (or be perceived as) more intuitive and simpler to use due to the direct and gesture-based interactions they enable (e.g., Watson, Hancock, Mandryk, & Birk, 2013). For SRL, which by definition is an active, effortful process (Schunk & Greene, 2018), this is particularly important in reference to perceived effort requirements. As previously outlined, learning with tablets has been associated with lower perceived expected effort (Mulet et al., 2019). Particularly for complex materials, that require deep effortful learning strategies to master (e.g., Azevedo, 2005), the assumption that little effort is required to learn the material can be inaccurate and potentially detrimental for learning. Further, perceptions of ease of use and effort requirements can inflate metacognitive judgments (Schwab et al., 2018), which in turn may lead to erroneous decisions regarding learning activities and suboptimal learning outcomes (e.g., Koriat, 2012). In these cases, additional self-regulatory processes need to be deployed to counteract the potential negative effect of inaccurate perceived effort requirements. On the other hand, tablets have great potential to foster SR through beneficial effects of gestures and hand proximity on attention or working memory load (Abrams et al., 2015; Goldin-Meadow et al., 2001). For instance, enhanced attention on relevant characteristics of the learning task can critically inform metacognitive judgments (Sidi, Shpigelman, Zalmanov, & Ackerman, 2017). More available cognitive resources can further be used to regulate learning processes (Follmer & Sperling, 2016). These examples showcase that the impact of tablet use for learning still remains unclear. From a SR perspective, multiple potential avenues for beneficial and detrimental effects of touch-interaction on SR processes and learning can be derived from the literature. Yet, little research has directly investigated the impact of tablets on SRL processes, particularly from a broader perspective on SR in education. Initial investigations on the interplay between tablet learning and SRL have shown that SR fosters learning of factual knowledge acquisition (Lee, 2015) and that self-reported SR skills enable successful learning with interactive tablet applications (Lee & Lee, 2018). It remains unclear from these studies, however, whether the same conclusion would hold for other aspects of SR. Therefore, based on previous systematic investigations of psychological correlates of academic achievement (Richardson et al., 2012), we propose to extend the investigation of SR when learning with tablets to include (1) learning strategies, (2) affective and motivational processes, (3) cognitive resources, and (4) personal dispositions.

3.1.5 The current study

In this paper we will analyze the impact of tablets (compared to PCs) on learning processes through the lens of SR. We aim to address this issue in greater detail by investigating if and how good a broad range of different self-regulatory constructs can predict learning outcomes in a complex and interactive learning environment presented either on a PC or a tablet. In particular, we aim at identifying the most predictive self-regulatory constructs for learning on PCs and tablets. Accordingly, we will focus on the following research questions.

1. Do learning outcomes significantly differ between participants who completed a learning task on a tablet and participants who completed the same task on a PC?
2. How accurately can learning outcomes be classified and does the accuracy significantly differ between tablet and PC?
3. Which self-regulatory constructs are the most robust and important predictors of learning outcomes when learning on PC or tablet and do these constructs differ between both types of learning environments?
4. How accurately can learning outcomes be classified by student’s self-regulatory dispositions, skills and strategies?

3.2 Data and Methods

3.2.1 Participants and procedure

Context of the experiment. The present study was part of an extensive experimental series on SR in educational contexts aiming at integrating the four different research traditions introduced in Section 1. To this end, 321 undergraduate students (age: M = 23.26 years, SD = 3.02 years, sex: 218 female, 95 male, 4 not specified) from a large, public German university completed 25 self-report surveys, 13 cognitive tasks, and five learning tasks related to SR in education. Up to fifteen participants were tested in parallel in three 4-hour sessions (including breaks) that took place on the same day of the week in three consecutive weeks. Participants were monetarily compensated or received course credits for participation and were awarded a bonus if they participated in all three sessions (total compensation up to 111€). The study was approved by a local ethics committee.

The art-learning task. The present study solely focused on the art learning tasks during the last session of the experimental series, in which learning with either a tablet or pc was experimentally manipulated. For this task students were randomly assigned to use a tablet or a PC to interact with a learning environment covering art historic contents. The learning environments on the two different devices were matched in physical size to ensure comparability between both conditions (327.66mm diagonal with a 4:3 aspect ratio in landscape orientation). Only participants with no missing data for any parts of this art learning task were included in our analyses, resulting in a final sample size of N = 291 students.

The art learning task lasted about 90 minutes and consisted of three parts. In a pre-learning phase, participants filled out questionnaires regarding their interest in art and related topics, and their current emotional experience before answering the 30-item art pretest. At the start of a 60-minute learning phase, participants were instructed to learn as much as possible about the contents of the learning environment (see below). However, no further guidance was provided on how to use the different parts of the learning environment for this purpose. Subsequently, in a post-learning phase, participants first had to indicate how confident they were in their ability to answer questions about the five topics covered in the learning environment (judgment of learning), before they completed a 30-item art posttest, which covered contents of the learning environment and was matched to the 30-item pretest. This exploratory task design with many degrees of freedom was chosen to enable students to self-regulate their learning substantially.

3.2.2 Materials

The learning environment. The learning environment was designed based on cognitive flexibility theory (Jacobson & Spiro, 1995) and in accordance with previous studies investigating the role of executive functions in multi-perspective hypermedia learning environments (MHEs, Kornmann, Kammerer, Zettler, Trautwein, & Gerjets, 2016). Specifically, the MHE consisted of two main components, an *artwork panel* and a *question panel* (see Figure 3.1). In the *artwork panel* an overview of 20 artworks from the Herzog Anton Ulrich-Museum in Braunschweig (Germany) was provided. Participants were able to rearrange the artworks according to their inventory number, a timeline or four art-historic perspectives such as country of origin or genre by touching the respective keywords at the bottom of the screen. Selecting an artwork enlarged it and enabled further interactions (e.g., moving and zooming). More importantly, when clicking an '*i*'-icon on the bottom corner of a selected artwork it was flipped over and an overview page appeared, on which a short summary paragraph about the artwork and five index cards with titles and short descriptions were displayed. Selecting an index card opened a content page with one to three paragraphs of text with accompanying pictures describing details about the artwork (e.g., explanations about the structure or technique).

In the *question panel*, participants had access to 12 guiding questions for each out of five relevant learning topics introduced. Participants were free to use them to guide their learning process as desired (e.g., by looking up the information required by the questions in the artwork panel and answering them). Furthermore, participants could use the *question panel* to display their remaining time in the learning session (by pressing a corresponding button) and to display their progress in the current topic (i.e., the percent of questioned completed) by means of a progress bar. Lastly, the question panel included a response confidence slider for each question that could be answered by participants as well as short self-reports at the end of each learning topic.

Both panels were integrated in a Qualtrics (Qualtrics, 2020) online survey. Participants were able to adjust the size of the panels in three steps (maximized *artwork panel*, default view, and enlarged *question panel*). However, the overall size of the app was restricted to full-screen size of the tablet in both conditions. The PC environment was controlled via mouse interaction, whereas the tablet was controlled via touch interaction. The controls of the learning environment were functionally equivalent in both conditions (i.e., in both conditions zooming was allowed: on the PC via the mouse wheel and on the tablet via pinch gestures).

Survey data. To explore the potential predictive value of different self-regulatory constructs for the art-learning task on PCs and tablets we selected self-report surveys that represent the traditions of research outlined in Section 1 (see Table

3.5 for a list of subscales used). These included self-reports capturing learning strategies (i.e., domain general learning and studying strategies, e.g., PISA approaches to learning scale Boerner, Seeber, Keller, & Beinborn, 2005; Schleicher, 1999), personality and other stable self-regulatory constructs (e.g., conscientiousness and grit Costa & McCrae, 2008; Duckworth et al., 2007), motivational measures (e.g., interest in art, expectancy value for art, motivational cost of dealing with art contents), general academic self-efficacy, and emotion regulation strategies (Abler & Kessler, 2009).

Cognitive task data. We selected measures from multiple cognitive tasks to represent key domains of working memory, cognitive abilities, and executive functioning (Table 3.6 for an overview of all cognitive task measures). Specifically, we included variables that represent the three executive functions updating (d-prime and inverse efficiency scores for a two- and three-back task,) switching (task switching accuracy and cost, Wisconsin card sorting task perseverance and total errors), and inhibition (stop signal reaction time, Stroop interference) as well as visuo-spatial abilities (Ekstrom, Dermen, & Harman, 1976) and working memory span (operation and reading span).

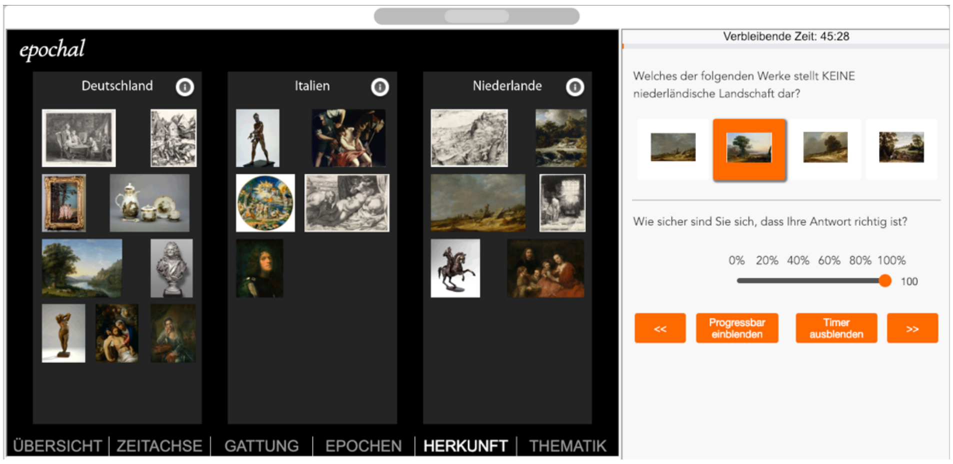


Figure 3.1. *Learning environment with artwork panel* (left; artworks sorted by origin) and question panel (right)

Learning outcomes. As a learning outcome measure a 30-item posttest covering the contents of the learning environment was designed by a domain expert.

This quiz consisted of 26 multiple-choice items (with four choice options) and four items, in which participants had to group four artworks into three categories or mark four areas of an artwork that depict specific artistic contents or styles. In preliminary analyses, we excluded items with a lack of variance in responses (i.e., less than 25% or more than 95% correct answers) and a negative correlation with average posttest score. According to these criteria six items were excluded, resulting in a 24-item posttest. Participants average performance in this posttest was used as learning outcome measure.

Groups for the classification models were obtained by splitting participants into the upper and lower 40% according to their average posttest performance (see Figure 3.2). We chose 40% as cutoff because this criterion yielded groups that were distinctive in posttest performance (upper 40%: M = 0.77, SD = 0.05, min = 0.70, max = 0.93; lower 40%: M = 0.51, SD = 0.07, min = 0.20, max = 0.59) while retaining a large sample size. Further, to test the stability of the classification outcomes and fluctuations in the importance of predictors in more extreme comparison, we split participants additionally in a separate analysis into upper and lower 30% in posttest performance.

Control variables. General mental ability (cultural fair intelligence test; Weiß, 2006), visuospatial ability (paper folding task; Ekstrom et al., 1976), reading ability (Schneider, Schlagmüller, & Ennemoser, 2007), and prior knowledge were included as control variables. Prior knowledge was obtained from a pretest that was structurally identical to the posttest (see 3.3.3 learning outcomes). 16 items of this test were maintained following the same preliminary analysis steps used for the learning outcomes.

Learning activity-related variables. To account for students' use of the learning environment, we further included three leaning-activity-related variables. These included the number of guiding items participants completed during the learning phase as well as the percent of these items that were answered correctly. Additionally, the mean value of students' judgements of learning completed after the learning phase was used as an indicator for their subjective confidence.

3.2.3 Analytical procedure

All analyses were conducted using the programming language Python (Van Rossum & Drake, 2011). Particularly, we used the ‘scipy’ module (Virtanen et al., 2020) for statistical tests and ‘scikit-learn’ toolbox for machine learning analyses (Pedregosa et al., 2011). For statistical tests, assumptions of normality and homoscedasticity were checked before analysis through visual inspection of histograms and Levene’s test (Levene, 1960). When violations of these assumptions were found, non-parametric tests were used instead (i.e., Welch’s test, Welch, 1947).

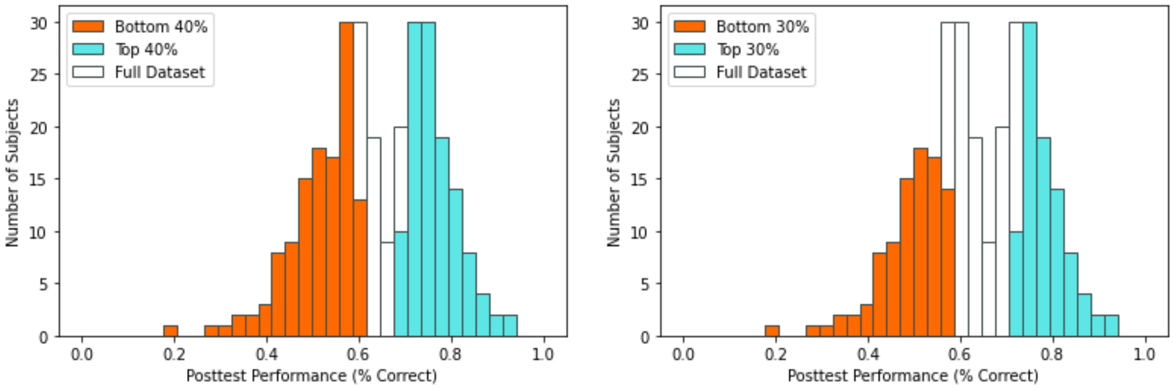


Figure 3.2. *Distribution of posttest scores by classification label*

Preliminary analyses. As preliminary analyses, all variables used as features for the classification and pretest scores were tested with regard to prior differences in mean, median, and variance between the Tablet and PC group using one-way ANOVAs, Kruskal-Wallis H-tests, Bartlett's tests and Levene's test. To ensure that predictions were not affected by innate differences in predictor variables, all analyses were separately conducted with and without variables that showed significant differences between the Tablet and PC group for any of the tests outlined above. The patterns of results were compared between both sets of analyses.

Classification. The classification consisted of four steps. (1) Features were standardized by calculating z-scores. (2) As a pre-selection step, false discovery rate corrected ANOVAs comparing the mean values of each predictor between the upper and lower 40% performers were conducted. Only features that showed significant differences at the five percent level after false discovery rate correction were selected for further analyses. (3) A cross-validated (5-fold) recursive feature elimination based on a l1-regularized logistic regression was used to eliminate unnecessary features among the features selected in the previous step. The area under the receiver-operator curve was used as performance measure in this step. (4) A support vector machine (SVM) was trained on the remaining features to predict whether a particular participant was in the upper or lower 40% according to posttest performance (or 30%, respectively).

We used a 10-fold cross validation to test the generalizability of these predictions. More specifically, steps (1) to (3) were applied to the training data set of each fold. More importantly, to identify the most robust and important self-regulatory variables and circumvent overfitting issues as much as possible, we decided to use ensemble predictions instead of a single optimized model. In detail, we repeated the 10-fold cross validation process 1000 times resulting in 10000 runs per medium and posttest value cutoff. The selected features and their importance for the prediction (i.e., the average decrease in model accuracy when a single feature is randomly permutated) were compiled for each run. Accuracies, were averaged across each 10-fold cross-validation. We applied this methodology to the entire data set (Tablet and PC conditions together) as well as the PC condition and the tablet condition separately.

Hyper-parameter selection. In line with the ensemble prediction method, we selected fixed hyper-parameters for our models. These parameters were obtained in initial tests using grid search using 10x10-fold nested cross-validation. Specifically, we examined different types of kernels (i.e., linear and radial basis function kernel) and a range of cost-parameters (ranging from 0.01 to 10). These preliminary tests revealed that a linear kernel was best suited for our data. Furthermore, cost parameters varied and showed an increase with greater numbers of selected features. Therefore, we opted to run all models with a linear kernel and a cost parameter equal to the number of features selected in each run.

3.3 Results

3.3.1 Preliminary analyses

Comparisons of mean values and standard deviations for alle predictors revealed that there were significant differences in mean/median values and/or standard deviations for four potential predictors between the experimental conditions. Elaboration strategies (inequal variances), the emotion regulation strategy cognitive reappraisal (mean values difference), agreeability (inequal variances) and the number of total errors in the Wisconsin card sorting task (inequal variances) showed potential issues (see Table 3.1). However, further analyses showed that removing these predictors from analyses had no impact on the pattern of results obtained. Therefore, in the following sections only results from models including all predictors will be reported, as no indication of a potential impact of the affected variables on the overall results were found.

Table 3.1.
Mean values and standard deviations for variables with innate differences between groups

Condition	Elaboration Strategies	Cognitive Reappraisal	Agreeability	Total Errors WCST
PC	2.14 (0.59)	4.71 (1.01)	119.96 (17.99)	0.17 (1.01)
Tablet	2.01 (0.73)	4.36 (1.13)	117.99 (21.39)	0.18 (1.13)

3.3.2 Performance across conditions (RQ1)

To test if there were significant differences in learning between Tablet and PC, a one-way ANOVA comparing the average posttest performance between Tablet and PC was conducted. Results showed no significant overall differences between the two conditions [$F(1,295) = 0.284$, $p = .594$, $\eta_p^2 = .001$]. Further, separate one-way ANOVAs were conducted for the upper and lower 30% and 40% separately to test whether the upper and lower performers differed between Tablet and PC. These analyses yielded no significant differences between Tablet and PC for the upper 40% [$F(1,117) = 0.208$, $p = .649$, $\eta_p^2 = .002$] and upper 30% in posttest performance [$F(1,87) = 0.012$, $p = .914$, $\eta_p^2 = .000$] as well as for the lower 40% [$F(1,119) = 0.964$, $p = .328$, $\eta_p^2 = .008$] and the lower 30% [$F(1,90) = 0.639$, $p = .426$, $\eta_p^2 = .007$]. This indicated that Tablet and PC conditions did not significantly differ in their overall patterns of learning outcomes.

3.3.3 Prediction Accuracy (RQ2)

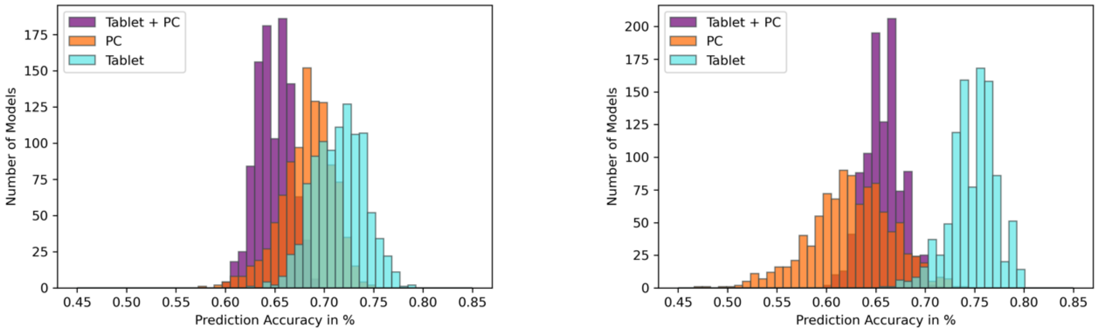
40% Cutoff. For classification of the upper versus lower 40% of performance in the art posttest we evaluated the accuracy across all classification models for Tablet and PC conditions combined as well as separately for the PC and the tablet condition. Initial investigations of the models showed that no feature met the criteria for pre-selection in 4.14% of runs of the PC models. For these iterations, no accuracies or feature importance could be computed. Therefore, the following results were obtained from only 9568 runs for the PC models, instead of 10000 runs for the tablet models as well as the models for Tablet and PC combined.

Across conditions (Tablet and PC combined), we found that upper and lower performing learners could be classified with 64.88% accuracy (*SD* = 1.58%, min = 60.00%, max = 68.75%). A permutation and a t-test comparing the accuracy of these models to the accuracy of the same models predicting randomly shuffled labels showed that the mean accuracy of our models was significantly higher than random predictions (*ps* < .001).

The mean classification accuracy in the PC condition was 68.16% (*SD* = 2.56%, min = 57.29%, max = 74.42%). Subsequent permutation and t-tests showed that the mean accuracy was significantly higher than the accuracy of models with randomized labels (*ps* < .001).

In the tablet condition, models averaged a prediction accuracy of 71.47% (*SD* = 2.46% min = 62.58%, max = 78.94%.). Equal to the aforementioned predictions, permutation and t-tests showed that the accuracy of these models was significantly higher than the accuracy of random predictions (*ps* < .001).

Lastly, the accuracy of the three types of models (Tablet and PC combined, PC, and tablet) were compared using t-tests. Results showed that tablet models where significantly more accurate than PC models (Welch's $t(1998) = 29.36$, $p < .001$), which in turn where significantly more accurate than models for tablet and PC combined (Welch's $t(1998) = 34.47$, $p < .001$).



Upper vs Lower 40%   Upper vs Lower 30%

Figure 3.3. *Distribution of classification accuracy for the Tablet + PC, PC, and tablet models. Classification accuracies were averaged for each 10-fold cross validation*

30% Cutoff. Models trained and tested on the upper versus lower 30% of learners showed that no feature met the feature selection criterions in 16.53% of the runs for the PC models. This resulted in 8347 runs for the PC models compared to the 10000 runs for both the tablet models as well as the models for Tablet and PC combined. Prediction accuracies for Tablet and PC combined were 65.72% on average (*SD* = 1.84%, min = 60.20%, max = 71.17%). PC models averaged an accuracy of 62.55% (*SD* = 4.13%, min = 47.22%, max = 74.11%), while tablet models had a mean accuracy of 74.84% (*SD* = 2.32%, min = 65.97%, max = 80.00%). Permutation and t-tests comparing each of these models to models with randomly permutated outcomes showed that all models performed significantly above chance level (*ps* < .001).

Finally, t-tests comparing accuracies of the three types of models showed that tablet models were significantly more accurate than models for Tablet and PC combined (Welch's $t(1998) = 97.21$, $p < .001$). Further, models for Tablet and PC combined were significantly more accurate than PC models (Welch's $t(1998) = 22.11$, $p < .001$).

3.3.4 Feature selection (RQ3)

In order to investigate which features were most important to differentiate upper and lower learners (based on their performance) overall and in each experimental condition, the percentage of runs in which the respective features were selected in the feature selection steps and their relative importance to the prediction were analyzed. The ten most important features for the 40% and 30% cutoff are displayed in Tables 2 and 3 and their mean values and standard deviations are shown in Tables 7 and 8 (Appendix C and D).

40% cutoff. Investigations of the most relevant features revealed distinct patterns for each type of model. In the combined models for Tablet and PC seven features were consistently used in the classifications. These predictors included prior knowledge, judgements of learning, critical evaluation, interest, reading comprehension, openness, and switching. Analyses of permutation importance further revealed that these features were similarly important with drops in accuracy for randomly permutating a variable ranging from 0.20% for interest to 1.64% for prior knowledge. Moreover, none of the less frequently selected features showed a positive impact on prediction accuracies.

Predictions on PC on the other hand were driven by two consistently selected features – prior knowledge and reading comprehension. Permutation importance showed that both of these variables had a large impact on the accuracy of these models with drops in accuracy of 8.92% for prior knowledge and 9.94% for reading comprehension. Analyses of less frequently selected features showed that the motivational value component (3.07%), interest (2.43%), and judgments of learning (0.29%) displayed limited positive contributions to prediction accuracy.

Tablet models showed two sets of predictors with regards to frequency. Critical evaluation, prior knowledge, and judgments of learning were consistently selected in the feature selection steps. Additionally, were motivational cost, openness, and switching frequently used to classify the upper and lower 40% with regards to posttest performance. Investigations of feature importance showed that critical evaluation had the largest impact on prediction accuracy (8.48%), followed by prior knowledge (5.29%), motivational cost (4.40%), and switching (3.05%). Judgments of learning (1.03%) and openness (0.83%) on the other hand had less impact on prediction accuracies. Of the less frequently selected features general mental ability (0.76%) and interest (0.29%) also showed small positive contributions to the prediction accuracy.

Table 3.2.

The ten most frequently selected features (40% cutoff)

Feature	Selected in % of Models			Importance ($\Delta_{Accuracy}$)		
	Combined	PC	Tablet	Combined	PC	Tablet
Prior Knowledge	100.00	98.89	99.54	1.64	8.92	5.29
Judgment of Learning	98.43	29.50	92.74	0.50	0.29	1.03
Critical Evaluation	99.65	0.05	99.95	0.44	-3.60	8.48
Interest	95.86	52.62	47.63	0.20	2.43	0.29
Reading Comprehension	95.99	99.61	0.00	0.81	9.94	-
Openness	96.61	0.90	86.74	0.23	-3.36	0.83
Switching	97.36	0.00	83.02	0.52	-	3.05
Cost	77.01	7.17	88.84	-0.14	-1.06	4.40
General Mental Ability	46.93	0.06	70.62	-0.40	-3.61	0.76
Motivational Value	31.06	63.80	20.00	-0.26	3.07	-0.62

30% cutoff. In the more extreme comparisons with regard to performance, the most commonly selected features for combined models for Tablet and PC remained largely unchanged. Most notably, reading comprehension and openness were less consistently included in models (reading comprehension: 63.27% instead of 95.99% and openness: 71.85% instead of 96.61%). Further, permutation importance showed that these variables did not positively contribute to prediction accuracy anymore when selected (reading comprehension: -0.89% and openness: -0.45%). Prior knowledge and judgments of learning were the most common and most important features in these models, followed by critical evaluation, switching, and interest. These features showed all positive contributions to the prediction accuracy ranging from 0.57% for interest to 3.10% for prior knowledge. Moreover, instrumental motivation was regularly included in models (selected in 88.51%) but on average showed no positive contribution (-0.53%) to the prediction when selected.

Similar to the 40% cutoff reading comprehension also remained a defining factor for the model accuracy (accuracy loss through random permutation: 7.32%). Moreover, prior knowledge (72.77%) and interest (78.58%) were less frequently selected and showed sizable impacts on classification accuracy (prior knowledge:

4.92% and interest: 2.09%). Lastly, the motivational value component was sparingly selected (69.02%) and also showed small positive effects (0.61%) on accuracy when included in the models.

For predictions on tablet, judgments of learning, switching, prior knowledge, and critical evaluation were consistently included in the feature selection processes. Similar to the 40% cutoff models, critical evaluation showed the largest impact on prediction accuracy (7.11%), followed by prior knowledge (6.76%) and switching (5.19%). While judgements of learning were more important in these models (4.09%), interest was less frequently selected (58.86%) and had little impact on prediction accuracies (0.16%). Lastly, visuospatial ability was sparingly selected (62.80%) and showed small positive effects (0.21%) on accuracy when included in the models.

Table 3.3.

The ten most frequently selected features (30% cutoff)

Feature	Selected in % of Models			Importance ($\Delta_{Accuracy}$)		
	Combined	PC	Tablet	Combined	PC	Tablet
Prior Knowledge	99.99	72.77	98.81	3.10	4.92	6.76
Judgment of Learning	99.93	55.97	99.82	1.95	-0.40	4.09
Interest	96.70	78.58	58.86	0.57	2.09	0.16
Switching	99.17	0.00	99.50	0.78	-	5.19
Critical Evaluation	98.77	0.00	98.75	0.87	-	7.11
Reading Comprehension	63.27	91.70	0.00	-0.89	7.32	-
Instrumental Motivation	88.51	53.59	0.19	-0.53	-0.96	-1.99
Visuospatial Ability	52.62	0.00	62.80	-0.60	-	0.21
Openness	71.85	0.38	37.07	-0.45	-5.70	-0.38
Motivational Value	18.74	69.02	4.56	-1.02	0.61	-3.43

3.3.5 Predictions without control variables

Prediction models without control variables (prior knowledge, reading ability, general mental ability, and visuospatial ability) or measures obtained during the learning phase (judgments of learning, number of items answered during the learning phase, number) were ran to get an indication to which extent self-regulatory variables exclusively can be used to classify upper and lower performers. Accuracies of these models showed a decrease in prediction accuracies for all models (see Table 3.4). Notably, the steepest decline in prediction accuracy was observed for PC models, in which the 40% cutoff model without control variables showed prediction accuracies at chance level. T-tests showed, that all other models still showed a significant drop in prediction accuracies when control variables were omitted (Welch's $t(1998) \leq -8.85$, $p < .001$), but still performed significantly above chance level (Welch's $t(1998) \leq -18.91$, $p < .001$).

Table 3.4.

Prediction accuracies with and without control and learning-phase-related variables

Model	With Control Variables		Without Control Variables	
	40% cutoff	30% cutoff	40% cutoff	30% cutoff
PC + Tablet combined	64.88 (1.58)	65.72 (1.84)	65.74 (1.67)	64.97 (1.94)
PC	68.16 (2.56)	62.55 (4.13)	49.82 (4.96)	55.38 (4.93)
Tablet	71.47 (2.46)	74.84 (2.32)	69.52 (2.95)	68.31 (2.53)

3.4 Discussion

This study investigated if and how learning outcomes, self-regulation and their interplay differ when learning with PCs and tablets. First, our results showed no differences in overall learning outcomes between these two media, which shows that successful learning is possible on both, PCs as well as tablets. This finding is in line with our expectations as previous research has shown a mixed picture, with potential advantages and disadvantages for learning with tablets (Mulet et al., 2019; Norman & Furnes, 2016; Sidi et al., 2017). More importantly, we investigated which self-regulatory requirements learning on PCs and tablets imposed onto learners. Taken together, our models revealed a very clear, yet surprising picture that learning on tablets demanded higher self-regulatory abilities from participants than learning on PCs.

Specifically, models predicting learning outcomes on PC were the least accurate and heavily dependent on control variables (i.e., reading comprehension and prior knowledge). This indicated that self-regulation during art learning on PCs was less important than the task-related predispositions that learners brought to the task. The important role of prior knowledge in self-regulated learning activities has been previously shown (Taub, Azevedo, Bouchet, & Khosravifar, 2014). However, even though previous studies have shown that reading comprehension is potentially more important on digital mediums, as they are negatively affected in digital contexts (Delgado, Vargas, Ackerman, & Salmerón, 2018), the predominant role of reading comprehension to differentiate upper and lower performers using PCs is unexpected. Especially given the complex multi-media stimuli used in the present study. Tablet models on the other hand were the most accurate models and incorporated a broad range of self-regulatory constructs, including learning activities (critical evaluation), driving forces (motivational cost), personal dispositions (openness), and cognitive resources (switching). Of these constructs critical evaluation was the most important predictor of learning outcomes on tablets. Specifically, high learning outcomes were achieved by students who evaluate materials more critically, which typically requires more effort investment than the surface level strategies (e.g., just reading or rehearsing the material). This indicates that the increased self-regulatory demands are driven by altered effort demands when learning with tablets (Mulet et al., 2019; Salomon, 1984; Schwab et al., 2018), that may result in shallower levels of processing. This explanation is further substantiated by the finding that upper and lower performers did not differ in the number of items they solved during the learning phase. In other words, low performers worked on the same amount of content on average, but did so with less critical reflection which resulted in suboptimal learning outcomes. The importance of critical reflection was even stronger than prior knowledge's impact on performance. Further, high learning outcomes on tablets required higher switching ability and openness as well as lower associated motivational cost with art learning compared to low learning outcomes on tablets. This demonstrates that a multitude of self-regulatory processes are required to overcome the altered perception of the learning task on tablets. Thus, opposed to previous research that indicated that tablets may foster learning by enhancing attention and cognitive processes (Abrams et al., 2015) or positively effect students motivation (Mulet et al., 2019), our results showed that tablets might even be detrimental for learning for students with deficits in the abovementioned self-regulatory processes.

Even though the set of self-regulatory variables investigated in the present study was extensive and the modelling approach used in this study was robust, the generalization of our findings to other content domains still requires further testing. The particular variables that we identified as most predictive in our research are potentially specific to this kind of learning task and environment. For instance, switching ability as cognitive resource was particularly potent due to the design of the learning environment that included two panels between which participants needed to integrate information. More linear tasks or other learning environments may impose different cognitive requirements. However, we argue while the specific predictors might differ, that variables representing the four facets of self-regulation we have been focusing on (learning strategies, affective and motivational processes, personal dispositions, and cognitive resources) may jointly predict learning best, particularly on tablets. Another limitation of this work and thus, a potential avenue for further research is that our modelling approach did not consider interaction effects explicitly. For instance, there is empirical evidence suggesting that invested effort and corresponding academic performance are determined by an interaction of conscientiousness and interest, rather than by their individual effects (Trautwein et al., 2015). As the prediction accuracies, even for the (best-predicting) tablet models, still indicate substantial room for improvement in accuracy, the approach of adding interaction terms as additional variables might lead to further improvements.

In sum, this study has shown that using a broader perspective on SR – including SRL, motivational constructs, cognitive resources, and personal dispositions - revealed differential self-regulatory requirements on PCs and tablets. Successful learning with tablets had higher self-regulatory requirements than learning on PCs. Future research can build on these results by applying this perspective on SRL to other learning situations and by further investigating the effects of learning with touch devices in other domains and applied settings. Ultimately, the design, development and use of touch technologies should be informed by all of the theoretical perspectives we have outlined to ensure that their potential to foster learning is fully exploited.

References

Abler, B., & Kessler, H. (2009). Emotion regulation questionnaire–Eine deutschsprachige Fassung des ERQ von Gross und John. *Diagnostica*, *55*(3), 144–152.

Abrams, R. A., Weidler, B. J., & Suh, J. (2015). Embodied Seeing: The space near the hands. In B. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 63, pp. 141-172). Amsterdam, The Netherlands: Elsevier. doi: https://doi.org/10.1016/bs.plm.2015.03.005

Agostinho, S., Tindall-Ford, S., Ginns, P., Howard, S. J., Leahy, W., & Paas, F. (2015). Giving learning a helping hand: finger tracing of temperature graphs on an iPad. *Educational Psychology Review*, *27*(3), 427–443.

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*(1), 20–29.

Azevedo, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? The role of self-regulated learning. *Educational Psychologist*, *40*(4), 199–209.

Bauer, I., & Baumeister, R. (2011). Self-regulatory strength. In K. Vohs, & R. Baumeister (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 64–82). NY: Guilford.

Bergmann, C., & Eder, F. (2005). *AIST-R. Allgemeiner-Interessen-Struktur-Test mit Umwelt-Struktur-Test (UST-R). Revision*. Göttingen, Germany: Hogrefe.

Bidjerano, T., & Dai, D. Y. (2007). The relationship between the big-five model of personality and self-regulated learning strategies. *Learning and Individual Differences*, *17*(1), 69–81.

Bissett, P. G., & Logan, G. D. (2011). Balancing cognitive demands: Control adjustments in the stop-signal paradigm. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *37*(2), 392–404. doi: 10.1037/a0021800

Boekaerts, M., & Corno, L. (2005). Self-Regulation in the Classroom: A Perspective on Assessment and Intervention. *Applied Psychology: An International Review*, *54*(2), 199–231.

Boekaerts, M., & Niemivirta, M. (2000). Self-regulated learning: Finding a balance between learning goals and ego-protective goals. In M.Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 417–450). New York: Academic Press. doi: 10.1016/B978-012109890-2/50042-1

Boekaerts, M., & Pekrun, P. R. (2015). Emotion and emotion regulation in academic settings. In L. Corno & E. M. Anderman (Eds.), *Handbook of educational psychology* (pp.76–90). New York, NY: Routledge.

Boerner, S., Seeber, G., Keller, H., & Beinborn, P. (2005). Lernstrategien und lernerfolg im studium. *Zeitschrift Für Entwicklungspsychologie Und Pädagogische Psychologie*, *37*(1), 17–26.

Brucker, B., Brömme, R., Ehrmann, A., Edelmann, J., & Gerjets, P. (2021). Touching digital objects directly on multi-touch devices fosters learning about visual contents. *Computers in Human Behavior*, *119*, 106708. doi: 10.1016/j.chb.2021.106708

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, *65*(3), 245–281. doi: 10.2307/1170684

Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior*. New York, NY: Cambridge University Press. doi: 10.1017/CBO9781139174794

Chamorro-Premuzic, T., & Furnham, A. (2003). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of Research in Personality*, *37*(4), 319–338.

Chu, M., & Kita, S. (2011). The nature of gestures' beneficial role in spatial problem solving. *Journal of Experimental Psychology: General*, *140*, 102–116. doi: 10.1037/a0021790

Cook, S. W. (2018). Enhancing learning with hand gestures: potential mechanisms. In K. D. Federmeier (Eds.), *Psychology of learningand motivation - advances in research and theory* (pp. 107–133). Amsterdam, NL: Elsevier. doi: 10.1016/bs.plm.2018.10.001.

Costa, P. T., & McCrae, R. R. (1998). Six approaches to the explication of facet-level traits: examples from conscientiousness. *European Journal of Personality*, *12*(2), 117–134.

Costa, P. T., & McCrae, R. R. (2008). The Revised NEO Personality Inventory (NEO-PI-R). In G. J. Boyle, G. Matthews & D. H. Saklofske (Eds), *The SAGE handbook of personality theory and assessment, Vol 2: Personality measurement and testing.* (pp. 179–198). Thousand Oaks, CA, US: Sage Publications, Inc. doi: 10.4135/9781849200479.n9

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, *26*, 197–223. doi: 10.1007/s10648-013-9246-y

Davoli, C. C., Du, F., Montana, J., Garverick, S., & Abrams, R. A. (2010). When meaning matters, look but don't touch: The effects of posture on reading. *Memory & Cognition*, *38*(5), 555–562.

de Bruin, A. B. H., & van Merriënboer, J. J. G. (2017). Bridging cognitive load and self-regulated learning research: A complementary approach to contemporary issues in educational research. *Learning and Instruction*, *51*, 1–9.

De Raad, B., & Schouwenburg, H. C. (1996). Personality in learning and education: A review. *European Journal of Personality*, *10*(5), 303–336. doi: 10.1002/(SICI)1099-0984(199612)10:5<303::AID-PER262>3.0.CO;2-2

Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. *Educational Research Review*, *25*, 23–38.

Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review*, *28*(3), 425–474. doi: 10.1007/s10648-015-9320-8

Dignath, C., Buettner, G., & Langfeldt, H.-P. (2008). How can primary school students learn self-regulated learning strategies most effectively?: A meta-analysis on self-regulation training programmes. *Educational Research Review*, *3*(2), 101–129.

Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, *3*(3), 231–264.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, *92*, 1087–1101. doi: 10.1037/0022-3514.92.6.1087

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*(1), 109–132.

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist*, *46*(1), 6–25. doi: 10.1080/00461520.2011.538645

Eilam, B., Zeidner, M., & Aharon, I. (2009). Student conscientiousness, self-regulated learning, and science achievement: An explorative field study. *Psychology in the Schools*, *46*(5), 420–432.

Eisenberg, I. W., Bissett, P. G., Enkavi, A. Z., Li, J., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Uncovering the structure of self-regulation through data-driven ontology discovery. *Nature Communications*, 10(1), 1-13.

Ekstrom, R. B., Dermen, D., & Harman, H. H. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.

Foglia, L., & Wilson, R. A. (2013). Embodied cognition. *WIREs Cognitive Science*, *4*(3), 319–325. doi: 10.1002/wcs.1226

Follmer, D. J., & Sperling, R. A. (2016). The mediating role of metacognition in the relationship between executive function and self-regulated learning. *British Journal of Educational Psychology*, *86*(4), 559–575.

Ginns, P., Hu, F., Byrne, E., & Bobis, J. (2016). Learning by tracing worked examples. *Applied Cognitive Psychology*, *30*(2), 160–169.

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science*, *12*(6), 516–522.

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, *38*, pp. 404–411. doi: 10.1037/h0059831

Gross, J. J. (2013). Emotion regulation: Taking stock and moving forward. *Emotion*, *13*, 359–365. doi: 10.1037/a0032135

Haßler, B., Major, L., & Hennessy, S. (2016). Tablet use in schools: A critical review of the evidence for learning outcomes. *Journal of Computer Assisted Learning*, *32*(2), 139–156.

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

Hofmann, W., Friese, M., Schmeichel, B. J., & Baddeley, A. D. (2011). Working memory and self-regulation. In K. D. Vohs, & R. F. Baumeister (Eds.), *Handbook of self-regulation: Research, theory, and applications (2nd ed.)* (pp. 204-225). New York, NY US: Guilford Press.

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

Hu, F.-T., Ginns, P., & Bobis, J. (2015). Getting the point: Tracing worked examples enhances learning. *Learning and Instruction*, *35*, 85–93.

Ilkowska, M., & Engle, R. (2010). Trait and state differences in working memory capacity. In A. Gruszka, G. Matthews, & B. Szymura (Eds.), *Handbook of individual differences in cognition: Attention, memory, and executive control* (pp. 295–320). New York, NY: Springer. doi:10.1007/978-1-4419-1210-7_18

Jacobson, M. J., & Spiro, R. J. (1995). Hypertext learning environments, cognitive flexibility, and the transfer of complex knowledge: An empirical investigation. *Journal of Educational Computing Research*, *12*(4), 301–333.

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*, 1-20.

Kang, S., & Tversky, B. (2016). From hands to minds: Gestures promote understanding. *Cognitive Research: Principles and Implications*, *1*(1), 4. doi: 10.1186/s41235-016-0004-9

Koriat, A. (2012). The relationships between monitoring, regulation and performance. *Learning and Instruction*, *22*(4), 296–298.

Kornmann, J., Kammerer, Y., Zettler, I., Trautwein, U., & Gerjets, P. (2016). Hypermedia exploration stimulates multiperspective reasoning in elementary school children with high working memory capacity: A tablet computer study. *Learning and Individual Differences*, *51*, 273–283. doi: 10.1016/j.lindif.2016.08.041

Kunter, M., Schümer, G., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., . . . Weiß, M. (2002). *PISA 2000: Dokumentation der Erhebungsinstrumente* [PISA 2000: Documentation of Scales]. Berlin, Germany: Heenemann GmbH.

Lee, H. W. (2015). Does Touch-based Interaction in Learning with Interactive Images Improve Students' Learning? *The Asia-Pacific Education Researcher*, *24*(4), 731–735.

Lee, H. Y., & Lee, H. W. (2018). The effects of cross-modality and level of self-regulated learning on knowledge acquisition with smartpads. *Educational Technology Research and Development*, *66*(2), 247–265.

Levene, H. (1960). Contributions to probability and statistics. *Essays in Honor of Harold Hotelling*, 278–292.

Levens, S. M., & Gotlib, I. H. (2012). The effects of optimism and pessimism on updating emotional information in working memory. *Cognition & Emotion*, *26*(2), 341–350.

MacLeod, C. M. (2005). The Stroop task in cognitive research. In A.Wenzel & D. C. Rubin (Eds.), *Cognitive methods and their application to clinical research* (pp. 17–40). Washington, DC: American Psychological Association. doi: 10.1037/10870-002

Marsh, H. W. (1990). A multidimensional, hierarchical model of self-concept: Theoretical and empirical justification. *Educational Psychology Review*, *2*(2), 77–172. doi: 10.1007/BF01322177

Mischel, W., Ayduk, O. N., Berman, M., Casey, B. J., Jonides, J., Kross, E., … Shoda, Y. (2011). Willpower" over the life span: Decomposing impulse control. *Social Cognitive Affective Neuroscience*, *6*, 252–256.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100.

Morell, M., Yang, J. S., Gladstone, J. R., Turci Faust, L., Ponnock, A. R., Lim, H. J., & Wigfield, A. (2020). Grit: The long and short of it. *Journal of Educational Psychology*. Advance online publication. doi: 10.1037/edu0000594

Mulet, J., Van De Leemput, C., & Amadieu, F. (2019). A critical literature review of perceptions of tablets for learning in primary and secondary schools. *Educational Psychology Review*, *31*(3), 631–662.

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–27). Cambridge, MA: MIT.

Norman, E., & Furnes, B. (2016). The relationship between metacognitive experiences and learning: Is there a difference between digital and non-digital study media? *Computers in Human Behavior*, *54*, 301–309.

OECD. (2013). *Trends Shaping Education 2013*. Paris, France: OECD. doi: 10.1787/trends_edu-2013-en

Ostendorf, F., & Angleitner, A. (2003). *NEO-Persönlichkeitsinventar nach Costa und McCrae, Revidierte Fassung (NEO-PI-R). Manual*. Göttingen, Germany: Hogrefe.

Oviatt, S., & Cohen, P. R. (2015). The paradigm shift to multimodality in contemporary computer interfaces. *Synthesis Lectures on Human-Centered Informatics*, *8*(3), 1–243.

Paas, F., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. *Educational Psychology Review*, *24*(1), 27–45. doi: 10.1007/s10648-011-9179-2

Panadero, E. (2016). A Review of Self-regulated Learning: Six Models and Four Directions for Research. *Frontiers in Psychology*, *8*, 1-28. doi: 10.3389/fpsyg.2017.00422

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, *12*, 2825–2830.

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

Pellegrino, J. W., & Hilton, M. L. (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. Washington, DC: National Academy of Sciences.

Pintrich, P. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). San Diego, CA: Academic Press. doi: 10.1016/B978-012109890-2/50043-3

Pintrich, P. R. (2003). A Motivational Science Perspective on the Role of Student Motivation in Learning and Teaching Contexts. *Journal of Educational Psychology*, *95*(4), 667–686. doi: 10.1037/0022-0663.95.4.667

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*, 322–338. doi: 10.1037/a0014996

Qualtrics. (2020). *Qualtrics*. Provo, Utah, USA: Qualtrics.

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.

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, *2*(4), 313–345.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, *132*, 1–25. doi: 10.1037/0033-2909.132.1.1

Rothbart, M. K. (1989). Temperament and development. In G. A. Kohnstamm, J. E. Bates, & M. K. Rothbart (Eds.), *Temperament in childhood* (pp. 187–247). New York: Wiley.

Salomon, G. (1984). Television is" easy" and print is" tough": The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology*, *76*(4), 647.

Schleicher, A. (1999). *Measuring student knowledge and skills: A new framework for assessment*. Paris, France: OECD.

Schneider, W., Schlagmüller, M., & Ennemoser, M. (2017). *LGVT 5-12+: Lesegeschwindigkeits-und-verständnistest für die Klassen 5-12: Manual*. Göttingen: Hogrefe.

Schunk, D. H. & Greene, J. A. (Eds.) (2018). *Handbook of Self-Regulation of Learning and Performance* (2nd Ed.). New York, NY: Routledge. doi: 10.4324/9781315697048

Schwab, F., Hennighausen, C., Adler, D. C., & Carolus, A. (2018). Television is still "easy" and print is still "tough"? More than 30 years of research on the amount of invested mental effort. *Frontiers in Psychology*, *9*, 1098. doi: 10.3389/fpsyg.2018.01098

Sepp, S., Agostinho, S., Tindall-Ford, S., & Paas, F. (2020). Gesture-based learning with ICT: Recent developments, opportunities and considerations. In S. Sepp, S. Agostinho, S. Tindall-Ford, & F. Paas (Eds.), *Advances in Cognitive Load Theory: Rethinking teaching*. New York., NY: Routledge.

Sidi, Y., Shpigelman, M., Zalmanov, H., & Ackerman, R. (2017). Understanding metacognitive inferiority on screen by exposing cues for depth of processing. *Learning and Instruction*, *51*, 61–73.

Sitzmann, T., & Ely, K. (2011). A meta-analysis of self-regulated learning in work-related training and educational attainment: What we know and where we need to go. *Psychological Bulletin*, *137*, 421–442. doi: 10.1037/a0022777

Sudevan, P., & Taylor, D. A. (1987). The cuing and priming of cognitive operations. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 89–103. doi: 10.1037/0096-1523.13.1.89

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, *72*(2), 271–324.

Taub, M., Azevedo, R., Bouchet, F., & Khosravifar, B. (2014). Can the use of cognitive and metacognitive self-regulated learning strategies be predicted by learners' levels of prior knowledge in hypermedia-learning environments? *Computers in Human Behavior*, *39*, 356–367.

World Bank. (2011). *Learning for all: Investing in people's knowledge and skills to promote development: Education sector strategy 2020*. Washington, DC: The World Bank.

Trautwein, U., Lüdtke, O., Nagy, N., Lenski, A., Niggli, A., & Schnyder, I. (2015). Using individual interest and conscientiousness to predict academic effort: Additive, synergistic, or compensatory effects? *Journal of Personality and Social Psychology*, *109*, 142–162. doi: 10.1037/pspp0000034

Tseng, P., & Bridgeman, B. (2011). Improved change detection with nearby hands. *Experimental Brain Research*, *209*(2), 257–269.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154.

Turvey, K., & Pachler, N. (2018). Tablet devices in education - beyond face value. In R. Luckin (Ed.), *Enhancing Learning and Teaching with Technology: What the research says*. London: UCL IOE Press.

Van Rossum, G., & Drake, F. L. (2011). *The python language reference manual*. Bristol, UK: Network Theory Ltd.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., … Bright, J. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272.

Watson, D., Hancock, M., Mandryk, R. L., & Birk, M. (2013). Deconstructing the touch experience. *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces*, 199–208.

Weidler, B. J., & Abrams, R. A. (2014). Enhanced cognitive control near the hands. *Psychonomic Bulletin & Review*, *21*(2), 462–469.

Weiß, R. H. (2006). *CFT 20-R: Grundintelligenztest skala 2-revision*. Göttingen, Germany: Hogrefe.

Welch, B. L. (1947). The generalization of student's' problem when several different population variances are involved. *Biometrika*, *34*(1/2), 28–35.

Wild, K.-P. and Schiefele, A. (1994). Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens [Learning strategies at university: Findings on the factorial structure and reliability of a new questionnaire]. *Zeitschrift fuer Differentielle und Diagnostische Psychologie*,*15*, 185/120.

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

Winne, P. H., & Hadwin, A. (2008). The weave of motivation and self-regulated learning. In D. Schunk & B. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 297–314). Mahwah, NJ: Erlbaum.

Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist*, *47*(4), 302–314.

Zheng, L. (2016). The effectiveness of self-regulated learning scaffolds on academic performance in computer-based learning environments: A meta-analysis. *Asia Pacific Education Review*, *17*(2), 187–202.

Zheng, R. Z. (2017). *Cognitive load measurement and application a theoretical framework for meaningful research and practice*. London: Routledge.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts & P. R. Pintrich, (Eds.), *Handbook of self-regulation* (pp. 13–39). San Diego, CA: Academic Press. doi: 10.1016/B978-012109890-2/50031-7

Appendix 3A

Table 3.5.

Self-report measures used in the present study

Questionnaire	Subscales	Items	α	References
Brief Self-Control Scale	-	13	0.83	(Tangney, Baumeister, & Boone, 2004)
Cross-Curricular Competencies	Control Strategies	5	0.58	(Kunter et al., 2002)
	Elaboration Strategies	4	0.82	
	Effort and Persistence in Learning	4	0.76	
	Self-Efficacy	4	0.77	
	Instrumental Motivation	3	0.91	
Emotion Regulation Questionnaire	Cognitive Reappraisal	6	0.81	(Abler & Kessler, 2009)
	Expressive Suppression	4	0.75	
General Interest Structure Questionnaire	Artistic Interest	10	0.84	(Bergmann & Eder, 2005)
GRIT	Consistency of Interest	7	0.79	(Morell et al., 2020)
	Perseverance of Effort	7	0.81	
Inventory for the Measurement of Learning Strategies in Academic Studies	Organization	9	0.67	(Boerner et al., 2005; Wild & Schiefele, 1994)
	Relationships	8	0.84	
	Critical Evaluation	8	0.88	
	Rehearsal	8	0.81	
	Effort	8	0.81	
	Attention	6	0.93	
	Time Management	4	0.87	
	Learning Environment	6	0.83	
	Learning with fellow students	4	0.80	
	Literature	4	0.80	
	Metacognitive Strategies	20	0.86	
Neo Personality Inventory	Agreeableness	48	0.90	(Costa & McCrae, 2008; Ostendorf & Angleitner, 2004)
	Conscientiousness	48	0.92	
	Extraversion	48	0.90	
	Neuroticism	48	0.93	
	Openness	48	0.86	
Subject Specific Motivation (Self-Concept, Interest, Value)	Self-Concept Art	4	0.89	
	Interest Art	2	0.93	
	Value Art	4	0.91	

Appendix 3B

Table 3.6.

Cognitive tasks and corresponding measures used in the present study

Task/Paradigm	Measure	References
CFT	Sum of correct responses	Weiß, 2006
N-Back Task	d-prime score (d')	Levens & Gotlib, 2012
Operation Span Task	Maximum span reached in adaptive O-Span task	Turner & Engle, 1989
PFT	Sum of correct responses	Ekstrom et al., 1976
Reading Comprehension Test	Comprehension Accuracy	Schneider, Schlagmüller, & Ennemoser, 2017
Reading Span Task	Maximum span reached in adaptive R-Span task	Turner & Engle, 1989
Stop Signal Task	Stop Signal Reaction Time	Bissett & Logan, 2011
Stroop Task	Reaction time difference for Color Word Interference Print Color Interreference	MacLeod, 2005
Switching Task	Reaction time difference between switch and non-switch trials Difference in accuracy between switch and non-switch trials	Sudevan & Taylor, 1987
Wisconsin Card Sorting Task	Count of omission errors Total count of Errors	Grant & Berg, 1948

Appendix 3C

Table 3.7.

Mean values and standard deviations of the ten most frequently selected features by condition (40% cutoff)

Feature (40%)	Combined		PC		Tablet	
	Upper	Lower	Upper	Lower	Upper	Lower
Prior Knowledge	0.48	0.58	0.48	0.57	0.47	0.58
	(0.13)	(0.13)	(0.13)	(0.12)	(0.13)	(0.14)
Judgment of Learning	2.05	2.54	2.09	2.49	2.01	2.58
	(0.80)	(0.77)	(0.75)	(0.74)	(0.85)	(0.73)
Critical Evaluation	1.95	2.36	2.01	2.26	1.87	2.44
	(0.77)	(0.76)	(0.80)	(0.82)	(0.74)	(0.71)
Interest	2.42	3.04	2.52	3.12	2.29	2.98
	(1.02)	(0.97)	(1.02)	(0.90)	(1.01)	(1.02)
Reading Comprehension	32.22	35.51	32.03	37.35	32.45	33.90
	(7.83)	(7.90)	(8.10)	(6.99)	(7.57)	(8.35)
Openness	122.63	133.23	125.73	131.61	118.91	134.67
	(18.36)	(16.04)	(16.88)	(15.41)	(19.50)	(16.58)
Switching	-0.04	-0.03	-0.03	-0.02	-0.05	-0.03
	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)
Cost	2.10	1.71	1.98	1.65	2.25	1.77
	(0.79)	(0.76)	(0.08)	(0.64)	(0.75)	(0.85)
General Mental Ability	36.97	39.34	37.94	39.40	35.81	39.29
	(5.25)	(4.81)	(4.68)	(4.36)	(5.68)	(5.20)
Motivational Value	2.22	2.73	2.28	2.78	2.14	2.69
	(0.86)	(0.81)	(0.86)	(0.86)	(0.86)	(0.93)

Appendix D

Table 3.8.

Mean values and standard deviations of the ten most frequently selected features by condition (30% cutoff)

Feature (30%)	Combined		PC		Tablet	
	Upper	Lower	Upper	Lower	Upper	Lower
Prior Knowledge	0.48	0.58	0.49	0.57	0.47	0.60
	(0.13)	(0.14)	(0.12)	(0.13)	(0.13)	(0.14)
Judgment of Learning	2.03	2.62	2.09	2.57	1.97	2.67
	(0.80)	(0.69)	(0.73)	(0.73)	(0.88)	(0.65)
Interest	2.40	3.15	2.51	3.23	2.27	3.08
	(1.05)	(0.97)	(1.03)	(0.88)	(1.07)	(1.04)
Switching	-0.04	-0.02	-0.03	-0.02	-0.05	-0.03
	(0.04)	(0.02)	(0.04)	(0.02)	(0.04)	(0.03)
Critical Evaluation	1.95	2.36	2.03	2.26	1.86	2.45
	(0.77)	(0.80)	(0.78)	(0.86)	(0.77)	(0.74)
Reading Comprehension	31.67	35.07	31.64	37.34	31.71	33.00
	(7.86)	(8.31)	(8.46)	(6.88)	(7.21)	(9.01)
Instrumental Motivation	1.82	1.32	1.97	1.40	1.66	1.25
	(0.97)	(0.90)	(0.93)	(0.93)	(1.00)	(0.88)
Visuospatial Ability	5.65	6.46	5.65	6.05	5.65	6.85
	(2.04)	(1.87)	(2.13)	(2.01)	(1.95)	(1.66)
Openness	122.75	133.89	126.08	132.47	118.95	135.22
	(18.82)	(16.43)	(17.82)	(15.31)	(19.41)	(17.48)
Motivational Value	2.20	2.83	2.28	2.90	2.11	2.77 (0.
	(0.90)	(0.90)	(0.89)	(0.86)	(0.91)	0.7894)

4 Study III

Multiple negative emotions during learning with digital learning environments: Evidence on their detrimental effect on learning from two methodological approaches

Franz Wortha, Roger Azevedo, Michelle Taub, Susanne Naricss

Note. Article published in Frontiers in Psychology used with permission of Frontiers Media SA

Estimated contributions

Scientific ideas by the candidate (%)	Data generation by the candidate (%)	Analysis and Interpretation by the candidate (%)	Paper writing done by the candidate (%)
80	0	90	85

Abstract

Emotions are a core factor of learning. Studies have shown that multiple emotions are co-experienced during learning and have a significant impact on learning outcomes. The present study investigated the importance of multiple, co-occurring emotions during learning about human biology with MetaTutor, a hypermedia-based tutoring system. Person-centered as well as variable-centered approaches of cluster analyses were used to identify emotion clusters. The person-centered clustering analyses indicated three emotion profiles: a positive, negative and neutral profile. Students with a negative profile learned less than those with other profiles and also reported less usage of emotion regulation strategies. Emotion patterns identified through spectral co-clustering confirmed these results. Throughout the learning activity, emotions built a stable correlational structure of a positive, a negative, a neutral and a boredom emotion pattern. Positive emotion pattern scores before the learning activity and negative emotion pattern scores during the learning activity predicted learning, but not consistently. These results reveal the importance of negative emotions during learning with MetaTutor. Potential moderating factors and implications for the design and development of educational interventions that target emotions and emotion regulation with digital learning environments are discussed.

Introduction

Learning is a complex multi-faceted process that requires students to deploy, monitor, and regulate their cognitive, metacognitive, affective and motivational processes based on the learning environment and the learning task and goal (Azevedo, Taub, & Mudrick, 2018). Emotions play a central role in this context. They significantly impact and drive processes that are quintessential to learning, such as attention, perception, memory (Lewis, Haviland-Jones, & Barrett, 2008; Tyng, Amin, Saad, & Malik, 2017), and metacognition (Azevedo, Mudrick, Taub, & Wortha, 2017). Furthermore, a long tradition of research has shown that emotions are directly related to learning outcomes and academic achievement (Boekaerts & Pekrun, 2015). Even though initial investigations on emotions and learning has almost exclusively focused on the importance of anxiety in learning and test situations (Pekrun, Goetz, Titz, & Perry, 2002), research on emotions and learning has diverged into investigations of a broad variety of affective states and emotions in differing learning contexts (e.g., classroom settings, research with advanced learning technologies or informal learning settings; Azevedo, Mudrick, Taub, & Bradbury, 2019). These studies have demonstrated that many different emotions are commonly experienced in learning settings (e.g., boredom, confusion, or frustration; D'Mello, 2013) and they have a significant impact on students' performance (e.g., D'Mello, Lehman, Pekrun, & Graesser, 2014; Pekrun et al., 2002). However, some important aspects of emotional experiences still have not been extensively researched in learning contexts. For example, most of the research in this context, particularly research during learning with digital learning environments, focused on the importance of single discrete emotions or sets of discrete emotions using variable-centered approaches. Research investigating emotions in other contexts, on the other hand, has revealed that approaches that consider multiple emotions simultaneously show great promise (e.g., Fortunato & Goldblatt, 2006; Vansteenkiste, Sierens, Soenens, Luyckx, & Lens, 2009). Only a few studies have investigated the complexity of students' (co-occurring) emotional experiences during learning using person-centered approaches (Ganotice Jr, Datu, & King, 2016; Jarrell, Harley, Lajoie, & Naismith, 2017; Jarrell, Harley, & Lajoie, 2016; Robinson et al., 2017; Sinclair et al., 2018). These studies have found that groups of students who differ in their emotional experiences during learning in regard to multiple emotions (so called emotion profiles) also meaningfully differ in their learning outcomes and academic achievement. The goal of this study was to combine person- and variable-centered approaches to examining emotions during learning with a digital learning environment. We extended upon previous research by considering a broader range of emotion measures than previous studies (i.e., academic achievement emotions and learning-centered emotions), incorporating emotion regulation and temporal dynamics of emotions, and by substantiating person-centered analyses with a novel variable-centered approach.

4.1 Emotions during learning with digital learning environments

Emotions are an essential component of learning activities across settings. Students' emotional experiences when learning with technologies are diverse, have been investigated on the basis of several frameworks (D'Mello, 2013), and have been classified in various categories, including academic achievement emotions (Pekrun et al., 2002), epistemic or learning-centered emotions (D'Mello & Graesser, 2012; Pekrun, Vogl, Muis, & Sinatra, 2017), and basic emotions (Ekman, 1992; Ekman & Friesen, 1971). Pekrun's (Pekrun et al., 2002; Pekrun, 2006) academic achievement emotions approach distinguishes academic emotions differing in their valence (positive vs. negative) and the perceived level of control by the learner, including enjoyment (positive and high control), anxiety (negative and medium control), and hopelessness (negative and low control). Learning-centered emotions approaches (also referred to as cognitive affective states or epistemic emotions; D'Mello, 2012; Pekrun et al., 2017) focus on emotions that are directly related to knowledge-generating aspects of cognitive processes (e.g., overcoming impasses during learning), including boredom, confusion and frustration. According to Ekman (1992) six basic emotions can be distinguished across cultural contexts and reliably identified from facial expressions, including anger, happiness, and surprise. An extensive amount of research has shown that emotions significantly impact learning processes, outcomes, and academic achievement (Pekrun & Linnenbrink-Garcia, 2014). The majority of studies revealed that the way emotions impact learning and achievement is closely related to their valence. More specifically, positive emotions are positively, and negative emotions are negatively related to the learning process and learning outcomes (e.g., Pekrun et al., 2002; Pekrun & Linnenbrink-Garcia, 2012; Pekrun, Lichtenfeld, Marsh, Murayama, & Goetz, 2017). However, there is also evidence opposing this general pattern. For example, studies identified detrimental effects of positive emotions on the accuracy of metacognitive judgements creating an illusion of learning (Baumeister, Alquist, & Vohs, 2015). Negative emotions on the other hand were positively associated with learning when they triggered deep processing of contents and were resolved by the students in a timely manner (see below, e.g., D'Mello & Graesser, 2014). This state of research indicates that, despite the overall tendency of beneficial effects of positive emotions and detrimental effects of negative emotions, further factors need to be considered to predict and explain the effects of emotions during learning.

A particular branch of research investigates (self-regulated) learning processes when learning with digital learning environments (Gegenfurtner, Fryer, Järvelä, Harackiewicz, & Narciss, 2019), including hypermedia learning environments (e.g., Opfermann, Scheiter, Gerjets, & Schmeck, 2013), intelligent tutoring systems (e.g., Azevedo et al., 2016; Harley, Taub, Azevedo, & Bouchet, 2017), and game-based learning environments (e.g., Sabourin & Lester, 2014; Taub, Azevedo, Bradbury, Millar, & Lester, 2018). These learning technologies have been designed and implemented to foster student learning about specific topics and have been shown to meaningfully enhance learning (Zheng, 2016). Digital learning environments include specific affordances that are directly linked to students' emotions. For example, research has demonstrated that the design of digital learning environments (e.g., shapes and colors; Plass, Heidig, Hayward, Homer, & Um, 2014), their structure (e.g., complex, non-linear structure; Arguel, Lockyer, Kennedy, Lodge, & Pachman, 2019), and scaffolds incorporated in such systems (e.g., prompts and feedback by pedagogical agents; Harley et al., 2017) can impact students' emotions. More specifically, digital learning environments can elicit and alter emotional processes or assist the learner in regulating them and provide unique opportunities for research to investigate emotions in ways hardly achievable in other contexts. For instance, multi-channel trace data can be collected with digital learning environments to measure emotions with minimal interruptions to the learning process (e.g., through automated detection of facial expressions; Azevedo et al., 2019; D'Mello, 2018). The dynamics of affective states model is a prominent theoretical framework in this line of research that focuses on the dynamic unfolding of specific, learning-centered emotions (D'Mello &

Graesser, 2012)[3]. More specifically, D'Mello and Graesser (2012) posited that confusion is elicited by impasses encountered during complex learning processes. This confusion can be beneficial to learning when it can be resolved, and the impasse can be overcome. Prolonged experiences of confusion on the other hand is theorized to lead to frustration and eventually boredom, which ultimately lead to disengagement and poor learning outcomes. Given that digital learning environments challenge students with learning tasks that require to develop a deep understanding of science concepts, or a solution for complex problems, such impasses are particularly likely to occur when learning with these systems. D'Mello and Graesser (2014) found a positive relation between (partially) resolved confusion and learning in a problem-solving task and a scientific reasoning task in an intelligent tutoring system. Another study by Taub et al. (in press) furthermore showed that the experience of frustration was linked to higher accuracy in the use of cognitive learning strategies (i.e., note-taking) with MetaTutor. However, they did not find a significant relation between emotions and learning gain.

Other studies on emotions and learning with digital learning environments (e.g., intelligent tutoring systems and game-based learning environments) on the other hand found detrimental effects of negative emotions. Initial studies on the relation between emotions and learning in AutoTutor identified significant detrimental effects of boredom for learning (Craig, Graesser, Sullins, & Gholson, 2004; Graesser, Rus, D'Mello, & Jackson, 2008). Across three studies using different digital learning environments, Baker, D'Mello, Rodrigo, and Graesser (2010) found further support for these findings by showing that boredom was the most persistent emotion (i.e., students were unlikely to transition from boredom to another emotion), and that boredom was the only emotion to be associated with maladaptive behaviors (i.e., gaming the system). Sabourin and Lester (2014) identified a positive relation between positive emotions and learning gains. Furthermore, they observed a negative association of confusion and boredom with learning gains in a game-based learning

---

[3] The dynamics of affective states model and related research often refer to cognitive-affective states instead of emotions. For consistency, readability, and because the cognitive component of these states resembles the appraisal component of emotion theories (e.g., Moors, Ellsworth, Scherer, & Frijda, 2013) we will refer to them as emotions. However, we acknowledge arguments that these terms might not be interchangeable in all contexts.

environment. A study by Grafsgaard et al. (2014) revealed that indicators of facially-expressed frustration were negatively predictive of learning gain.

Taken together, these studies demonstrated the importance of learning-centered emotions during learning with digital learning environments (for a recent review see Arguel et al., 2019). However, they also demonstrated a profoundly controversial relation between (negative) emotions and learning. This clearly indicates further research is needed to disentangle the manifold relation between emotions, learning, and learning outcomes by identifying factors that explain these contradictory relations. One such factor that has been rarely considered in the aforementioned studies on emotions in digital learning environments is the co-occurrence of emotions. Even though studies have shown that the emotions outlined above have differential effects on learning depending on other affective states they are accompanied by or lead to (e.g., D'Mello & Graesser, 2012; Goetz et al., 2014; Riemer & Schrader, 2019), the co-occurrence of emotions and the breadth of emotional experiences has rarely been considered in this context.

4.2 Person centered approaches to emotions

Research on emotions during self-regulated learning has indicated that a variety of emotional states and processes impact learning and performance in meaningful ways. While these studies have greatly contributed to a comprehensive understanding of emotions in learning situations, especially when learning with digital learning environments, they have not fully considered the breadth of emotional experience of an individual. More specifically, the variable-centered approach used by these studies focuses on singular emotional states or a pre-selected set of emotions while controlling for the impact of other emotions. Emotion research on the other hand suggests that individuals can experience multiple emotions concurrently, and that these emotions affect each other reciprocally, which ultimately impacts thoughts and behaviors (e.g., Fernando, Kashima, & Laham, 2014; Lazarus, 2006). Person-centered approaches typically identify groups of students with similar emotional experiences in regard to multiple emotions at a certain point of time (often referred to as emotion profiles). These profiles are then compared to another and related to relevant outcome measures (e.g., learning and academic achievement). For example, multi-level investigations of affect in college students have revealed that spurs of negative emotions coupled with positive trait affectivity are associated with greater academic growth than positive or negative affect alone (Barker, Howard, Galambos, & Wrosch, 2016). Furthermore, the added value of this approach has been repetitively shown outside of educational contexts (e.g., Fernando et al., 2014; Vansteenkiste et al., 2009). In research in education settings, this approach is still quite rare. We identified five studies that used a person-centered analytical approach in different educational contexts (see Table 4.1 for a brief overview).

Jarrell and colleagues (2016; 2017) investigated emotions when learning with a computer-based learning environment using a person-centered approach in two studies. Five discrete emotional states (enjoyment, pride, hope, shame, and anger) measured with the Achievement Emotions Questionnaire (AEQ; Pekrun et al., 2002) were used to cluster students with similar emotional experiences. In both studies, a three-profile solution including a positive, negative, and low emotional experience profile, was identified. These profiles were subsequently related to learning outcomes. The first study (N = 26) revealed no significant differences in performance between profiles. In the follow-up study (N = 30) Jarrell et al. (2017) investigated differences in diagnostic performance efficiency between emotion profiles. They found that the negative emotion profile was outperformed by at least one other profile averaged across levels of difficulty (easy, medium, hard) and for easy and hard tasks, but not for tasks with medium difficulty.

Further investigations of emotions through a person-centered approach were conducted by Ganotice Jr et al. (2016) in two secondary school samples. Similar to the studies outlined above, discrete emotional states (enjoyment, hope, pride, anger, anxiety, shame, hopelessness, boredom) measured through the AEQ (Pekrun et al., 2002) were used for clustering. In a domain general or a math-specific context, four emotion profiles were identified. These profiles included a high positive and high shame profile, a moderate positive and negative emotion profile, a high negative emotion profile, and a high positive emotion profile. These profiles were compared in regard to school engagement, motivation, and math performance. Results showed that profiles with high positive emotions were the most adaptive profiles while the high negative emotion profile was the least adaptive.

Robinson et al. (2017) investigated affective profiles in an undergraduate anatomy course. Other than previous person-centered studies, this research used two dimensions of affect (positive/negative x activated/deactivated, see Ben-Eliyahu &

Linnenbrink-Garcia, 2013) as clustering variables. Through a two-step procedure, they identified four emotion profiles including a positive, a deactivated, a negative, and a moderate negative profile. Comparison in academic achievement revealed that the deactivated profile showed higher academic achievement than both negative profiles (negative and moderate negative) throughout three exams. Robinson et al. (2017) also found differences between the positive and the negative profile, but not throughout all exams. Lastly, they investigated the mediating role of (dis-) engagement and found that higher levels of performance for the positive and deactivated profile were mediated through lower levels of disengagement.

Lastly, Sinclair et al. (2018) investigated emotion profiles displayed in an undergraduate student sample that learned about the human circulatory system using MetaTutor (see 5.3 MetaTutor). They used five discrete emotion states (enjoyment, curiosity, pride, boredom, and frustration) measured at five time points before and during learning using latent profile analysis. Similar to the studies above, they found a positive, negative (bored/frustrated), and moderate emotion profile. Subsequently they investigated transitions between profiles and found that students from the negative profile were least likely to transition to another profile. Lastly, they found that learning gain predicted the transitions between profiles at specific, selected time points.

Taken together, these studies demonstrate that a person-centered approach can reveal emotion profiles across contexts, ranging from laboratory studies to research in schools and university. Furthermore, all studies have found that these profiles are significantly related to performance, academic achievement, and related constructs. Most of the previous studies have not incorporated learning-centered or epistemic emotions (e.g. boredom, confusion, and frustration; D'Mello & Graesser, 2012). On the other hand, previous research on emotions when learning with digital learning environments has found that these emotions significantly impact learning in varying ways. The finding that these emotions can have a positive or negative impact on learning is particularly interesting for person-centered research as the contradicting implications might be explained though co-occurring emotions (i.e., profiles that show similar levels of confusion or frustration, but varying levels of other emotions). The only study that investigated learning-centered emotions (Sinclair et al., 2018) on the other hand did not consider achievement emotions in their analysis, which makes comparisons across studies difficult. We aim to address this issue by including learning-centered emotions in addition to academic achievement emotions that were used in most of the person-centered studies outlined so far.

Furthermore, the aforementioned studies have investigated different constructs related to emotions and performance such as motivation (Ganotice Jr et al., 2016) or engagement (Robinson et al., 2017) to substantiate their findings. None of the studies investigated the role of emotion regulation in this context. Emotion regulation is an essential component to emotional experiences in learning contexts and is a critical link between emotional experience and academic outcomes (Gross, 2015; Harley, Pekrun, Taxer, & Gross, 2019). It describes students' efforts to influence which emotions they experience, when they experience these emotions and how they express them (Harley, Pekrun, Taxer, & Gross, 2019). Emotion regulation strategies are for example the cognitive reappraisal of emotional experiences or modification of the situation that elicited the emotion (Gross, 2015). Spann, Shute, Rahimi, and D'Mello (2019) found that emotion regulation significantly influenced the relation between emotions and learning in a game-based learning environment. More specifically, they found that cognitive reappraisal led to higher learning outcomes for highly confused, frustrated, and engaged students, but was not as effective for students with low levels of confusion, frustration and engagement. Incorporating emotion regulation could shed light on the development of emotions in relation to specific profiles. Adaptive profiles (such as described by Ganotice Jr et al., 2016) are potentially defined by higher levels of emotion regulation to cope with high levels of negative emotions. To investigate this subject matter, temporal investigations of emotions related to emotion profiles similar to Sinclair's approach (Sinclair et al., 2018) are necessary. This includes, investigating the self-reported use of emotion regulation strategies for the different emotion profiles, and exploring to what extent the intensity of emotional experiences fluctuates over time within profiles.

Lastly, the studies outlined above were limited to using person-centered approaches only. While the great value of this type of research has been shown, we argue that supplementing person-centered with other approaches can be essential to their understanding. More specifically, identifying if the distinguishing characteristics of profiles (e.g., varying levels of positive or negative emotion intensity) can be replicated through variable-centered approaches can provide additional insight on the origin of these profiles. Such approaches could differentiate if profiles are based on natural co-occurrence of emotions (e.g., high correlations between negative emotions) or specific combinations of individual emotional experiences (e.g., a profile with high levels of boredom and other negative emotions versus a profile with high levels of boredom and low levels of other negative emotions). Furthermore, replicating results using two different methodologies would reveal their level of robustness, which is particularly important in this context, because emotion profiles are identified through data driven approaches (guided by previous research).

4.3 Current Study

The current study aims to address the issues outlined above by identifying emotion profiles of students who learned with MetaTutor and relate them to learning outcomes. To this end we decided to adapt the person-centered analytical procedure outlined by Robinson et al. (2017) and Vansteenkiste et al. (2009) for the identification of emotion profiles. Additionally, we demonstrate how a variable-centered approach can substantiate these results by relating patterns of emotions to emotion profiles and learning outcomes throughout different phases of learning (i.e., before the learning phase, at the start of the learning phase, and at the end of the learning phase, see 5.4.1 Emotion items). More specifically, we aim to answer the following questions.

1.1 *Which emotion profiles can be identified during SRL with MetaTutor and how can they be described?* Given that the specific profiles are highly dependent on the number of clusters, no specific hypothesis can be formulated a priori. However, based on previous person-centered studies, we expect a negative and positive emotion profile (see Ganotice Jr et al., 2016; Jarrell et al., 2016, 2017; Robinson et al., 2017; Sinclair et al., 2018). Additionally, further likely profiles can include a low-intensity or moderate intensity profile for all emotions.

1.2 *Are there significant differences in learning outcomes between the profiles?* Based on previous research, we expect the profile with the highest values of negative emotions to display the lowest learning gain (Ganotice Jr et al., 2016; Jarrell et al.,2016, 2017; Robinson et al., 2017).

1.3 *Are there significant differences in self-reported use of habitual emotion regulation strategies between the profiles?* Based on research on emotion regulation, we expect profiles characterized by high negative emotion intensities to indicate lower levels of self-reported use of emotion regulation strategies (Harley, Pekrun, Taxer, & Gross, 2019).

2.1 *How can stable patterns of emotions can be identified throughout the different phases of the learning session and how can they be described?* Similar to our first research question, we expect a strong differentiation between negative and positive emotions in the different phases. Additionally, a strong differentiation between activating and deactivating emotions is expected (Ben-Eliyahu & Linnenbrink-Garcia, 2013). Furthermore, because neutral – per definition – refers to the absence of perceivable and detectable emotions, we hypothesize neutral to represent its own cluster (potentially paired with emotions that show low intensities overall). Lastly, based on the reoccurring finding that specific emotions are positively and/or negatively related to learning, we expect boredom, confusion or frustration to form separate cluster(s) from other negative emotions (e.g., D'Mello & Graesser, 2012).

2.2 *How are emotion profiles related to the phase-specific patterns of emotions?* We expect emotion profiles to significantly differ in regard to emotion clusters that are defined by valence as all previous studies included profiles that were defined by positive and negative emotions (Ganotice Jr et al., 2016; Jarrell et al.,2016, 2017; Robinson et al., 2017). In an exploratory step we will investigate if these differences are stable over time or if they arise throughout specific parts of the learning session.

2.3 *How can the phase-specific patterns of emotions predict learning outcomes in the respective phases of the learning activity?* Based on previous research, we expect negative emotions to be most predictive of learning. However, the direction of this interaction will be explored, as previous research has shown controversial results in this regard.

4.4 Methods

4.4.1 Participants

One hundred ninety-four (N = 194) undergraduate students (aged between 18 and 41, M = 20.46 years, SD = 2.96 years; 53% female) from three large public North American universities participated in a two-day laboratory study. They were randomly assigned either to the *prompt and feedback* (P+F) or *control* (C) condition (see 5.3. MetaTutor), and monetarily compensated for their time ($10 per hour, up to $40). For the present study, only participants that filled out a sufficient number of emotion questionnaires (see 5.4.1 Emotion items) were included in analyses, resulting in a sample size of one hundred seventy-six (*N* = 176) students.

4.4.2 Procedure

The experiment was conducted over two days. On the first day, participants signed a consent form, filled in demographics questions, and completed several self-report measures (e.g., the Achievement Emotions Questionnaire – Pekrun et al., 2002 and the Emotion Regulation Questionnaire – Gross & John, 2003). Lastly, after responding to the questionnaires, participants took a 30-item pretest about the human circulatory system.

On the second day of the experiment, students were first introduced to the learning task and learning environment. They were instructed to set two learning sub goals before the beginning of the learning phase. During the learning phase, participants had to engage in self-regulated learning by reading texts, inspecting corresponding diagrams, and completing quizzes. Moreover, regardless of the experimental condition (see section 5.3 MetaTutor) students were free to indicate their use of certain cognitive (e.g., note taking) or metacognitive learning strategies and activities using the SRL palette implemented in MetaTutor's interface (see 5.3 MetaTutor). Additionally, quizzes and self-report measures (i.e., the emotion and values [EV] questionnaire; Azevedo et al., 2013) were administered based on specific rules implemented by the system (e.g., the EV was conducted on a time-based threshold – roughly every 14 minutes during the learning session with MetaTutor).

After the 60-minute learning phase, students were directed to the post test (i.e., 30-item test about the circulatory system) and had to complete a last set of self-reports (e.g., an EV directly before the posttest) before they were debriefed by the research assistant.

During the experiment, multiple channels of multimodal data, including eye tracking, galvanic skin response, and automated analysis of facial expressions were collected. However, these process measures were not analyzed in the present study.

4.4.3 MetaTutor

MetaTutor is a hypermedia-based tutoring system that fosters self-regulated learning processes while learning about the human circulatory system (Azevedo et al., 2018). The system was designed using a set of production rules, which fire based on how students monitor and control their understanding of the text and relevancy of the current page to the sub-goal they are working on. In addition to the processes being prompted by the pedagogical agents based on the production rules, participants were able to engage in any process of their choice. The MetaTutor learning environment was strategically designed to foster the use of cognitive learning strategies and metacognitive monitoring processes (see Figure 3.1). For example, a timer (A) and sub goal progress bar (C) allow students to monitor their progress toward achieving their sub goals and overall learning goal. The table of contents (B) provides students all the content page titles so they can select the appropriate pages to read for achieving their sub goals. There are seven pre-set sub goals in the environment (path of blood flow, heartbeat, heart components, blood vessels, blood components, purposes of the circulatory system, and malfunctions of the circulatory system). Prior to the 60-minute learning session, students are progressed through a sub goal setting phase where they are guided to set two of those sub goals. The content text (D) and diagram (E) facilitate knowledge acquisition and foster coordinating information between the text and diagram. The SRL palette (F) provides students the opportunity to select cognitive learning strategies (i.e., prior knowledge activation, take notes, summarize, make an inference) and metacognitive monitoring processes (judgment of learning, feeling of knowing, content evaluation) they want to use during learning about the human circulatory system.

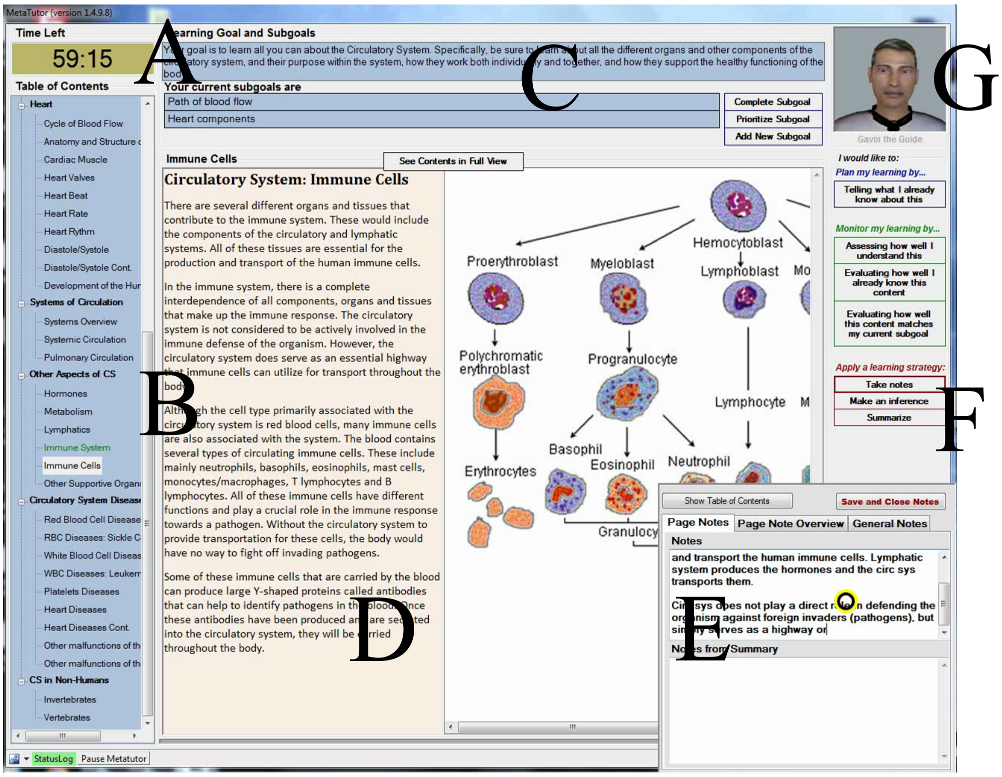


Figure 4.1. *Screenshot of the MetaTutor interface.*

There are four pedagogical agents with one present at a time (G), where each agent focuses on a specific component of SRL. Gavin (shown in Figure 4.1) guides students through the learning environment and administers self-report questionnaires. Pam fosters planning by helping students set sub goals and activate their prior knowledge. Sam focuses on strategy use. Mary emphasizes monitoring processes. The amount of assistance provided by the pedagogical agents depends on the experimental condition students are assigned to. In the P+F condition, the agents prompt students to engage in SRL processes (using time- and event-based production rules). They also provide feedback on how they performed. For example, Sam will prompt students to make a summary, and once they have done so, he will tell them it is too long, too short, acceptable, etc. In the C condition, students are not prompted by the agents, nor are they given any feedback on their performance. In this condition, students can still initiate the use of cognitive and metacognitive processes, however they are still not given any feedback, whereas in the P+F condition, students can also self-initiate the use of these processes, and will be given feedback on their performance.

4.4.4 Measures

4.4.4.1 Emotion items

Students' emotional experiences at the start, during, and at the end of the learning session were measured using the Emotion-Values Questionnaire (EV; Azevedo et al., 2013). The EV covers 15 emotional states as well as two questions asking about the perceived value and the students' ability to perform well on the current task on a five-point Likert-scale (ranging from 1 – "Strongly Disagree" to 5 – "Strongly Agree"). Additionally, two forced choice items asked the participants to select the emotion that best describes how they currently feel out of 15 (all emotional states from the EV) and 7 options (basic emotions), respectively. The emotional states included in the EV were based on extensive research on achievement emotions in academic settings (Pekrun, 2006; Pekrun et al., 2002), as well as on research on learning-centered emotions/epistemic emotions (e.g., D'Mello & Graesser, 2012; Pekrun et al., 2017). The questionnaire covers the following emotions (in order of administration): enjoyment, hope, pride, frustration, anxiety, shame, hopelessness, boredom, surprise, contempt, confusion, curiosity, sadness, eureka, and neutral. A definition and an example were provided for each emotional state during each administration.

The EV was administered at fixed points of time before and after the learning phase, and time-based during the learning phase. More specifically, the questionnaire was administered directly before and after participants set their learning sub goals, and before the actual learning phase. During the learning activity the questionnaire was administered every 14 minutes. Lastly, the final EV was administered when the learning phase was finished, directly before the post test. The number of EVs completed varied between participants because the administration during the learning phase was postponed when key learning activities took place. In particular, the questionnaire did not interrupt any of the user- or agent-initiated learning strategies that required completing quizzes or filling out questionnaires. For example, if a student initiated the sequence of finishing the current learning sub goal, they had to fill out a 10-item multiple-choice quiz on the current topic and received feedback depending on the experimental condition (see section 5.3). If an EV should have been administered during that sequence, it was postponed until the end of the sequence, potentially delaying it by several minutes. This resulted in a range of four to eight EVs completed between participants. To allow for comparisons of participants, we decided to limit the EVs analyzed in the present study to six points of time relative to the start and the end of the learning session. Therefore, only participants that completed at least six EVs were considered for analyses, resulting in a final sample size of one-hundred-seventy-six students (N = 176). The following EVs were selected: (1) the first two EVs that were completed at the beginning and end of the sub goal setting phase, (2) the third and fourth EV, which took place in the first half of the learning phase, and (3) the last two EVs, which was the last questionnaire presented during the learning phase, and the final EV immediately prior to the post test. Due to missing data, “Eureka” was excluded from analyses in the current study, yielding 14 discrete emotions considered for analyses.

4.4.4.2 Pre and post tests

Prior knowledge and learning outcomes were measured using two 30-item multiple choice tests covering conceptual knowledge of the human circulatory system. The measures were developed by a domain expert in the subject matter. Each question had four potential answers and one correct solution. The order of two equivalent versions of the tests was randomized and counterbalanced across experimental conditions. Percent correct for both measures were computed for analyses.

4.4.4.3 Emotion regulation questionnaire

Students’ self-reported habitual use of emotion regulation strategies was measured using the emotion regulation questionnaire (ERQ; Gross & John, 2003). The ten-item questionnaire features two sub-scales asking about the use of emotion regulation strategies using a seven-point Likert scale (ranging from 1 – strongly disagree to 7 – strongly agree). More specially, mean values for the sub-scales expressive suppression (4 items, α = .78; e.g., “I keep my emotions to myself.”) and cognitive reappraisal (6 items, α = .84; e.g., “I control my emotions by *changing the way I think* about the situation I’m in.”) were calculated for analyses.

4.5 Statistical Analyses

Statistical analyses in the present study were conducted using R (R Core Team, 2019), Python (Van Rossum & Drake, 2011), and SPSS (SPSS, 2012). Before the initial analyses, we investigated if the mean scores for each emotion computed over the six administrations of the EV for clustering contained significant outliers using Grubbs (1969) approach (implemented through the 'grubbs.test' function of the outlier package for R; Komsta, 2011). In total, twelve univariate outliers were replaced by the closest non-outlier value (three for shame, one for hopelessness, two for surprise, two for confusion, and five for surprise). Furthermore, investigations of the skewness and kurtosis (values < 2; George & Mallery, 1999) revealed that all of the variables used for analyses (i.e., mean emotion scores, emotion cluster scores and learning measures) were within acceptable ranges of normal distribution.

The person-centered methodological approach for the identification of emotion profiles was based on previous studies investigating affective, emotional, or motivational profiles (Robinson et al., 2017; Vansteenkiste, 2009). More specifically, we first used the 'hclust' function of R's stats package to compute a range of profile solutions using Ward's method and extracted the cluster centroids for each profile. We used agglomeration coefficients obtained through the SPSS classification function (SPSS, 2012), minimum number of profile size (Fernando et al., 2014), and cluster fit indices from 'Nbclust' (Charrad, Ghazzali, Boiteau, Niknafs, & Charrad, 2014) to identify the eligible range of clusters. Subsequently, k-means clustering analysis with these centroids as starting points was conducted ('kmeans' function of the 'stats' library) to obtain the most distinctive set of profiles. As a last step in the cluster identification, we used the cross-validation procedure outlined by Breckenridge (2000) to assess the stability of the solution (using self-implemented function based on the 'knn' function of the 'class' library; Venables & Ripley, 2002). Together with investigations of explained variance in the clustering variables and redundancy of the clusters, this criterion was used to determine the final cluster solution. Clustering methodology was chosen because the suitability of clustering over other methodological approaches in this context has been repeatedly showcased by previous research (e.g., Robinson et al., 2017).

Subsequently we used a latent growth linear mixed effect model to investigate differences in learning outcomes between emotion profiles. Models were fit using 'lmer' from the ‘lme4’ library (Bates, Mächler, Bolker, & Walker, 2014). Summary statistics were extracted via the ‘analyze’ function of ‘psycho’ (Makowski, 2018) and post-hoc comparisons were conducted using ‘glht’ from ‘multicomp’ (Hothorn, Bretz, & Westfall, 2008). Additionally, this analysis was repeated for all profile solutions (including the initial solutions from hierarchical clustering) to assess if the findings were stable throughout different profile configurations.

Then spectral co-clustering – a machine learning clustering approach – implemented through the ‘SpectralCoclustering’ function of the Python library ‘scikit-learn’ was used to substantiate the relation between emotions and learning identified through the profiling approach (Pedregosa et al., 2011). Specifically, we grouped emotions into clusters based on their correlation across all measurement points and separately for each time point. The emotion cluster solution was selected based on its stability over all administrations of the EV and alignment to previous research. Then, principal component analysis (‘PCA’ function of ‘scikit-learn’) with one main component was used to obtain participants’ scores for each emotion cluster at each measurement point. Additionally, the internal consistency of emotion clusters was assessed though Cronbach’s Alpha (‘alpha’ of R’s ‘psych’ package; Revelle, 2017). The obtained scores were then used in multiple regressions for each time point separately to assess how the emotion clusters are related to learning. Regression weights were calculated using the ‘lm.beta’ function R’s ‘lm.beta’ package (Behrendt, 2014).[4]

4.5.1 Preliminary Analyses

To control for the potential effect of the experimental manipulation of the present study (i.e., the control and prompt + feedback conditions) on the results described in the following sections, all variables included in the analyses were compared between the experimental conditions using multivariate analyses of variance (MANOVAs). Results showed no systematic differences in pre- and posttest scores, emotion scores, or emotion cluster scores between the conditions (all $p > .05$; except negative emotions cluster scores for EV 1: $p < .05$). Additionally, we conducted

[4] Analyses scripts and data are available upon request. For analyses that required (pseudo) randomization, seeds used to obtain the results reported in this paper were documented to guarantee replicability.

chi-square tests for each profile solution to test if the experimental conditions were equally represented in each emotion profile. Results revealed no significant differences in the distribution of experimental conditions for any of the emotion profiles identified.

4.5.2 Person-centered Approach: Emotion Profiles

4.5.2.1 Identifying emotion profiles

To identify emotion profiles, students with similar self-reported emotional experiences were grouped using a two-step clustering approach. More specifically, first, hierarchical clustering (Ward's method) was used on the squared Euclidian distance matrix for the mean values of each emotion for each participant throughout all six time points (see above). Each participant started as their own cluster in the hierarchical clustering analyses. Then the closest participants were merged into a cluster. This step was repeated until all participants were merged into a single cluster, resulting in a range of cluster solutions between the number of participants (i.e., each participant as their own cluster) and a singular cluster. To identify the profile solutions eligible for subsequent analyses, we used three criteria: (1) the scree-plot of agglomeration coefficients to identify the point where the addition of clusters did not substantially decrease the agglomeration coefficient, (2) a sufficient profile size for statistical analyses (n > 10; Fernando et al., 2014), and (3) multiple cluster fit indices (Charrad et al., 2014). Agglomeration coefficient indicated that merging a three-cluster solution into two clusters was not practical ($\Delta_{coefficient}$ = 233.93). A second drop in agglomeration coefficients was identified for the addition of a sixth cluster, but was less substantial ($\Delta_{coefficient}$ = 73.379). While this procedure favored solutions with more than six profiles, the second criterion limited the number of profiles to a maximum of seven, as all further profile solutions included profile(s) with less than ten participants. Lastly, we compared the solutions that were sufficient for both criteria in regard to 26 fit indices (see Charrad et al., 2014 for a complete list of the indices) and found equal support for the three to five profile solutions and little to no support for the six and seven profile solutions. Accordingly, the three-, four-, and five-profile solutions were selected for further analyses.[5] Preliminary analyses on the structure of the clusters revealed a noteworthy feature. A single emotion profile with higher negative emotion intensities than other profiles (n = 29) was a stable component of all solutions outlined above.

As a second step in the identification of emotion profiles we used k-means clustering, a non-hierarchical clustering procedure, in order to increase similarity within clusters and differences between clusters. More specifically, for the previously selected three- to five-cluster solutions, we first extracted the cluster centroids. These values were then used as starting points of the k-means clustering instead of starting with randomized seeds. In this procedure the number of clusters is defined a priori. Then a starting seed was used as the initial centroid of a cluster and participants that were in proximity to that centroid (measured through a distance threshold) were assigned to that cluster. This procedure was repeated for each starting seed until all participants were assigned to a cluster (Fortunato & Goldblatt, 2006). K-means clustering was chosen because this procedure simultaneously maximizes between cluster distances (i.e., increased differences between emotion profiles) and minimizes within-cluster variance (i.e., increased similarity within profiles; Eshghi, Haughton, Legrand, Skaletsky, & Woolford, 2011). After obtaining the respective cluster solution, we then assessed the rate of agreement between the hierarchical and k-means approaches. Both clustering methods showed sufficient rate agreement (K_3 = .76; K_4 = .78; K_5 = .78). This indicated that the k-means clustering altered the initial profiles obtained through the hierarchical clustering but maintained the overall structure and demonstrates the robustness of the identified profiles. To test if the aggregation of self-reported emotion intensities had a significant impact on the obtained emotion profiles, we re-ran all previous steps using all six measurement points for the fourteen emotions as clustering variables. Comparison of the profiles identified by clustering means and the profiles identified by clustering all measurement points demonstrated high to very high agreement (K_3 = .85; K_4 = .91; K_5 = .88). This indicated that our data supports the use of mean values as clustering variables and further underlined the robustness of the clustering procedure.

[5] All subsequent analyses were also conducted for the six- and seven profile-solution. The pattern of results remained similar. These results were not reported in this study due to space constraints.

To select the emotion profiles for subsequent analyses we first compared the explained variance in mean emotion intensities between the solutions with different numbers of emotion profiles. The three-profile solution explained moderate levels of variance for all mean emotion intensities, except neutral, surprise, anxiety and contempt (see Table 4.2). The four-profile solution explained more variance for most of the emotions, but also showed lower levels of explained variance for specific emotions (i.e., contempt and confusion). This pattern also applied to the comparison of the four- and five-profile solutions. However, while the four-profile solution added a profile that was primarily defined by boredom in addition to the neutral, positive, and negative emotion profiles of the three-profile solution, the five-profile solution only added a profile that was largely redundant to the positive emotion profile (with higher levels of curiosity, surprise and anxiety). Based on the largely redundant nature of this profile (a criteria used by Fernando, Kashima, & Laham, 2014), we decided not to consider this solution.

Table 4.2
Explained variance by profile-solution

Emotion	Profile solution		
	3	4	5
Enjoyment	0.47	0.63	0.62
Hope	0.46	0.55	0.54
Pride	0.34	0.39	0.40
Frustration	0.31	0.38	0.41
Anxiety	0.20	0.22	0.42
Shame	0.55	0.60	0.65
Hopelessness	0.64	0.67	0.68
Boredom	0.44	0.60	0.62
Surprise	0.14	0.23	0.39
Contempt	0.22	0.13	0.12
Confusion	0.40	0.34	0.43
Curiosity	0.36	0.49	0.46
Sadness	0.39	0.44	0.45
Neutral	0.10	0.15	0.23
Average	0.36	0.42	0.46

As the final step for selecting the most suitable cluster solution, we cross validated the three- and four-profile solutions following the procedure outlined by Breckenridge (2000). More specifically, we split our sample randomly into two equally large sub samples. Then, the two-step clustering procedure outlined above was separately applied to each of the sub samples. The two sub samples were subsequently compared with a k-nearest-neighbors approach. More specifically, each participant of a sub sample was assigned to a new cluster value based on their most similar counterparts in the other sub sample (their nearest neighbors). To assess the robustness, Kohen's Kappa (as a measure for agreement) was calculated based on the initial (obtained through the two-step approach) and new cluster assignment (obtained through the nearest neighbors procedure) in both samples. To increase the robustness of the cross-validation, we repeated this procedure twenty times and averaged Kappa values across all iterations (i.e., twenty-fold cross validation). Results indicated that the three-profile solution (K = .65) showed sufficient stability (i.e., K > .60; Asendorpf, Borkenau, Ostendorf, & Van Aken, 2001; Breckenridge, 2000), but the four-profile solution did not (K = .56). Therefore, the three-profile solution was selected as the final profile solution (see Figure 4.2 for a comparison of mean emotion intesities between the three profiles). Means and standard deviations for mean emotion intensities, and pre and post test scores of the three-profile solution are displayed in Table 4.3. The three profiles can be described by their most distinct features as follows[6]. The first profile (n = 75) displayed low to moderate levels for all emotions except boredom and neutral, which were at moderate levels. The neutral score was higher than for the other profiles. Accordingly, we refer to this profile as *neutral*. The second profile (n = 62) showed moderate to high levels for most of the positive emotions (joy, hope, pride, curiosity) and low levels of negative emotions (frustration, shame, hopeless, boredom, contempt, confusion, and sadness). The positive emotion intensities were higher in this profile compared to those of the other profiles. Thus, we labeled this profile as the *positive* emotion profile. The final profile (n = 39) was characterized by medium levels for all emotions. When compared to the other profiles, the most distinct feature of this group was their increased levels of negative emotion intensities for all negative emotions. Therefore, we referred to this group as the *negative* emotion profile. A multivariate analyses of variance (MANOVA) revealed that the emotion profiles significantly differed in regard to their mean emotion intensities (Wilks's λ (28, 320) = 0.100, p < 0.001, η^2 = 0.68).

4.5.2.2 Linking emotion profiles and learning outcomes

Differences in learning outcomes between profiles were analyzed using a latent growth linear mixed effect model. More specifically, we predicted learning outcomes with time (pre and post test) and profile membership as fixed factors and included a random intercept[7] based on previous studies that showed the importance of individual differences in prior knowledge when learning with MetaTutor (Taub, Azevedo, Bouchet, & Khosravifar, 2014). The model explained significant proportion of variance in

[6] Labels for the profiles were chosen based on the most dominant feature overall and in comparison to the other profiles.
[7] The step-wise model selection procedures were not reported due to space constraints. They can be found in the supplementary materials.

learning outcomes (R^2 = 68.03%; fixed effects: R^2 = 16.23%) and showed that learning outcomes significantly improved over time for all profiles (β = .75, SE = 0.06, t(175) = 12.36, p < .001, VIF = 1.00) and that membership in the negative profile was associated with significantly lower learning outcomes (β = -0.40, *SE* = 0.16; t(173) = -2.43, *p* < .05, *VIF* = 1.18; see Figure 4.3)[8]. Post hoc test using Tukey's HSD (honestly significant difference) showed that significant differences in learning outcomes were only found between the negative and neutral profile (*z* = -2.432; *p* < .05).

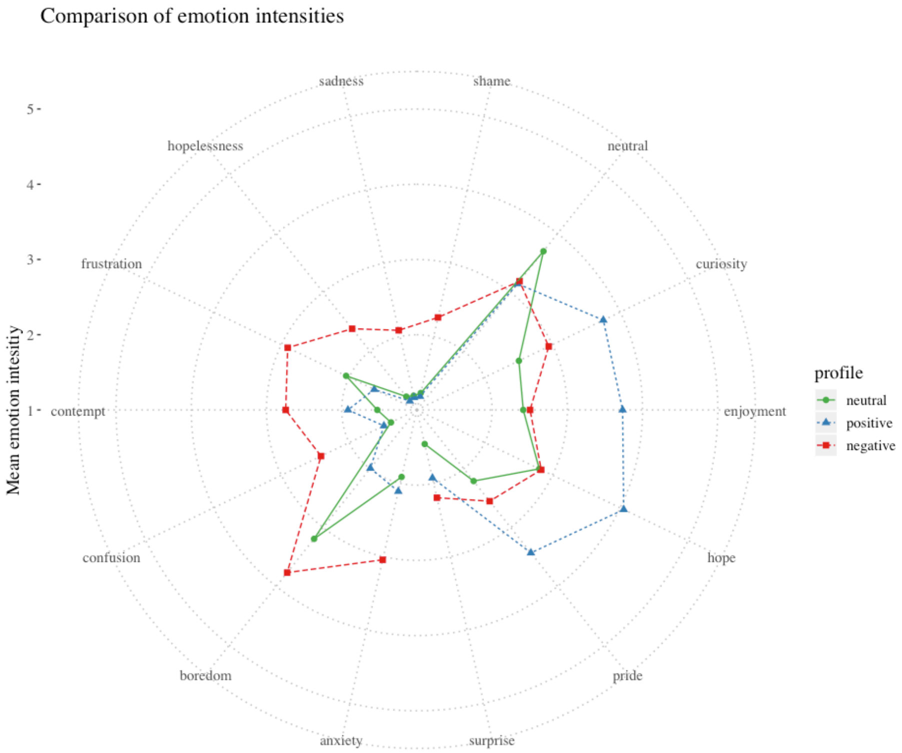


Figure 4.2. *Comparison of mean emotion intensities between profiles*

4.5.2.3 Linking emotion profiles and emotion regulation

Two separate ANOVAs comparing expressive suppression and cognitive reappraisal between the profiles were conducted to test if profiles differed in their self-reported habitual use of emotion regulation strategies. Results showed that there were

---

[8] This pattern of results was consistent for all profile solutions of the two-step approach but only for the three-profile solution in the initial clustering approach.

no significant differences in expressive suppression ($F(2, 163) = 0.013$; $p = .99$), but significant differences in cognitive reappraisal between profiles ($F(2, 163) = 4.185$; $p < .05$). Post-hoc comparisons using Bonferroni correction revealed that students with a negative emotion profile had significantly lower cognitive reappraisal scores ($M = 4.62$, $SD = 1.27$) then those with a positive emotion profile ($M = 5.30$, $SD = 1.07$; $p < .05$).[9]

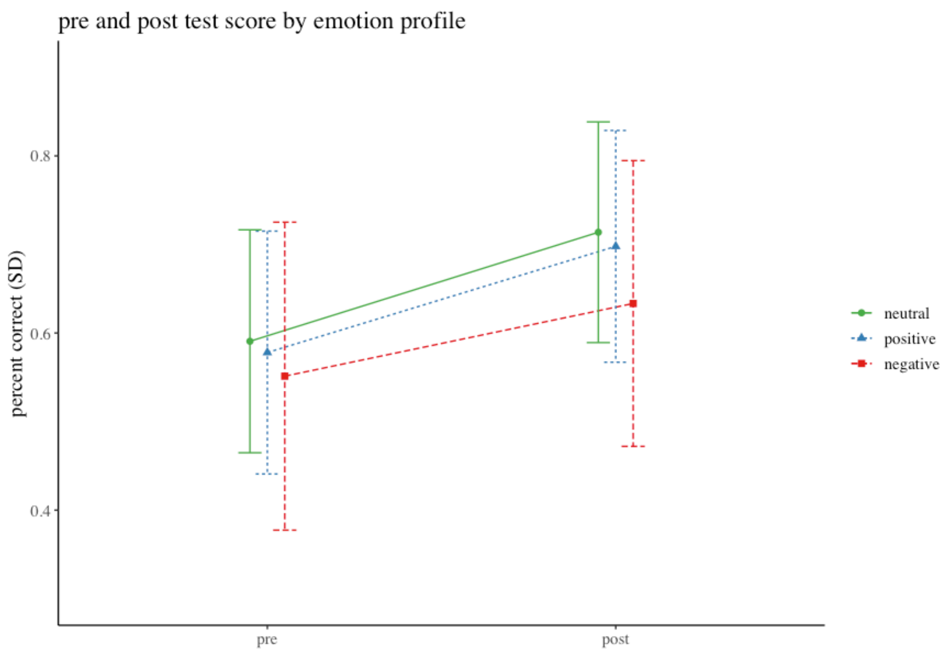


Figure 4.3. *Pre and post test scores by emotion profile*

4.5.3 Variable-centered Approach: Patterns of Co-occurring Emotions

4.5.3.1 Identifying patterns of co-occurring emotions

To identify patterns of co-occurring emotions, correlation matrices for the 14 emotions investigated in this study were computed separately for each point of time (see procedure) and aggregated over all points of time. Then, spectral co-clustering, a clustering technique that groups data by rows and columns simultaneously (e.g.,

[9] Further analyses revealed that this pattern of results stayed identical across all profile solutions (four- and five-profile solution from the k-means clustering as well as three- to five-profile solutions from the initial hierarchical clustering analyses).

Kluger, Basri, Chang, & Gerstein, 2003), was applied to these matrices to obtain the variable-centered patterns of related emotions for each point of time and aggregated over all EV administrations. This procedure was carried out for cluster solutions ranging from three to six clusters. A four-cluster solution was the only one that displayed great stability over all time points and aggregated over all measures (the only exception is that contempt moved to the boredom cluster during the last measurement). This solution included a positive and a negative emotions pattern, as well as neutral and boredom as singular-emotion clusters (see Table 4.4). Cronbach's Alpha was calculated for the negative and positive emotions pattern separately for each time point to test if the identified cluster represented an internally consistent linear structure sufficiently well. Results showed that both the negative pattern (alpha ranging from .74 to .81) and the positive pattern (alpha ranging from .72 to .85) met this criterion.[10] We obtained participants' individual scores for each pattern and the maintained variance of each pattern through principal component analyses with one component. The maintained variance from the original Likert-scale items for each non-singular emotion pattern was sufficient in this solution (35.45% for the negative emotions pattern for EV2 and 68.40% for the positive emotions pattern for EV2, see Table 4.4). Loadings for all emotions were positive for each pattern (i.e., increases in emotion intensity was associated with an increase in pattern score).

4.5.3.2 Exploring differences in variable-centered emotion patterns scores between emotion profiles

Differences in emotion variable-centered cluster scores between profiles over time were analyzed using latent growth linear mixed effect models. More specifically, we predicted variable-centered emotion pattern scores with time (six administrations of the EV), profile membership and their interaction as fixed factors and included a random intercept for the negative, positive and boredom emotion patterns.[11] The model for the neutral emotion pattern did not include the interaction term of time and profile membership as the addition of this factor did not improve the model significantly.

---

[10] No item had to be rescaled for these analyses, showing that all measures in the pattern correlated positively with the pattern score assigned to each participant.

[11] Random slopes were initially considered but lead to potentially overfitted models (singular fit) and were therefore not considered in final analyses. The model selection summary can be found in the supplementary materials.

Results showed significant differences in emotion pattern scores on average for all emotion clusters (all $p < .001$). Furthermore, the negative, positive, and boredom pattern scores showed significant linear growth for all participants (all $p < .001$). For negative emotion pattern scores ($R^2 = 62.99\%$, fixed effects: $R^2 = 40.54\%$) we found significantly different linear trajectories between the negative profile and the other profiles (compared to neutral profile: $\beta = 0.22$, $SE = 0.05$, $t(877) = 4.57$, $p < .001$, $VIF = 4.44$; compared to positive profile: $\beta = 0.21$, $SE = 0.05$; $t(877) = 4.16$, $p < .001$, $VIF = 4.11$; see Figure 4.4). Linear growth in positive emotion pattern scores ($R^2 = 66.10\%$, fixed effects: $R^2 = 40.44\%$) were significantly different between the positive and other profiles (compared to neutral profile: $\beta = 0.08$, $SE = 0.04$, $t(877) = 1.99$, p < .05, $VIF = 3.17$; compared to negative profile: $\beta = 0.17$, $SE = 0.05$, $t(877) = 3.43$, $p < .001$, $VIF = 2.59$). Boredom pattern scores ($R^2 = 56.99\%$, fixed effects: $R^2 = 25.58\%$) illustrated significantly different linear trajectories between the positive and the neutral profile ($\beta = 0.14$, $SE = 0.05$, $t(877) = 3.14$, $p < .010$, $VIF = 3.28$).

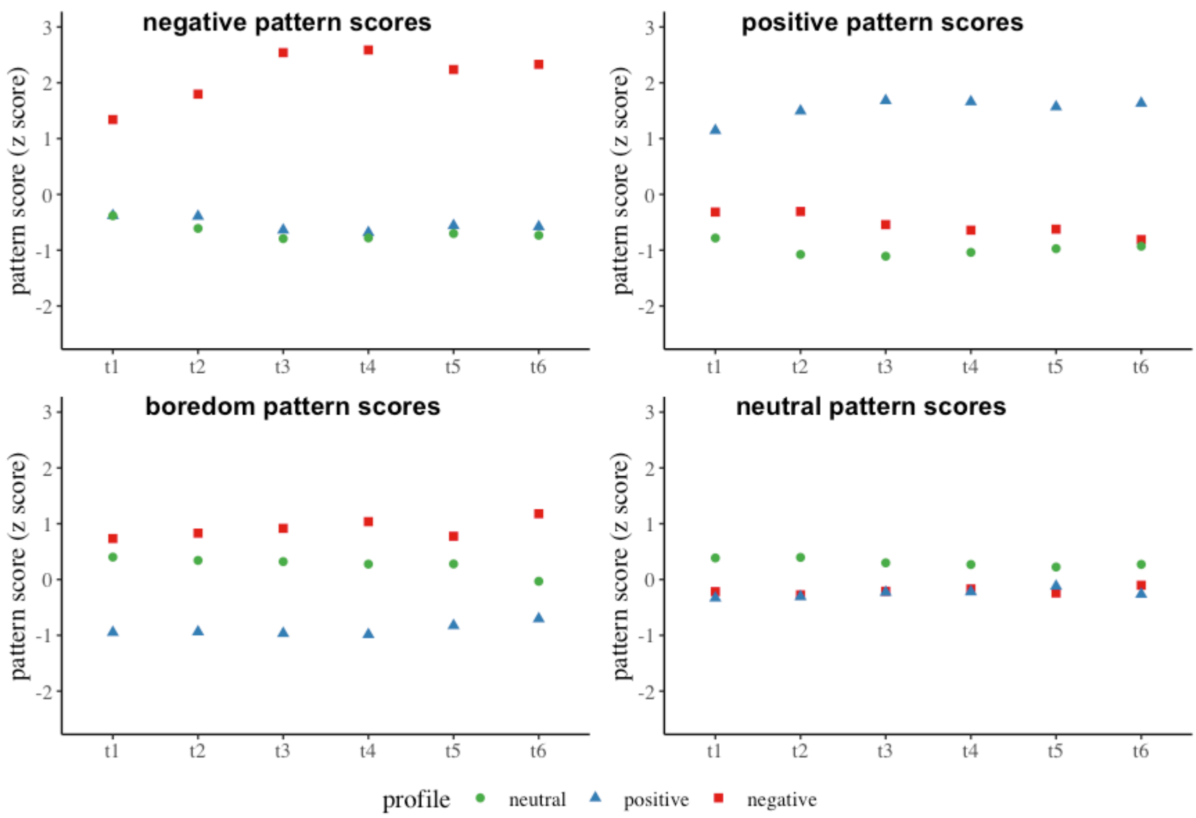


Figure 4.4. *Emotion pattern scores by emotion profile over the six measurement points*

4.5.3.3 Linking of co-occurring emotions and learning outcomes

To asses if variable-centered emotion patterns can predict learning gains, separate linear regression models predicting post test score with pretest and variable-centered emotion pattern scores for each point of time were calculated. Results showed that pretest score was a significant predictor of post score in all regressions (β ranging from .58 to .62; p < .01). The explanatory value of variable-centered emotion pattern scores beyond the effect of pretest score throughout the different points of time varied. The positive emotions pattern was the only significant predictor besides pretest score for the first administration of the EV (before learning sub goals were set; $F(5,170) = 26.03$, $R^2 = .42$; $\beta = .15$; $p < .05$) and a marginally significant predictor for the second administration (after learning sub goals were set; $F(5,170) = 25.30$, $R^2 = .41$; $\beta = .14$; $p = .057$). Negative emotions pattern scores significantly predicted post test score for the fourth (second EV during the learning activity; $F(5,170) = 24.22$, $R^2 = .40$; $\beta = -.13$; $p < .05$) and sixth administrations of the EV (directly before the post test; $F(5,170) = 25.08$, $R^2 = .41$; $\beta = -.17$; $p < .05$) and were a marginally significant predictor for the third (first EV during the actual learning activity; $F(5,170) = 24.05$, $R^2 = .40$; $\beta = -.11$; $p = .086$) and fifth EVs (last EV during the learning activity; $F(5,170) = 23.01$, $R^2 = .39$; $\beta = -.11$; $p = .082$). The other patterns showed no significant relation to post test score at any time point.

4.6 Discussion

This study used a person-centered approach to identify emotion profiles and a variable-centered approach to identify variable-centered emotion patterns throughout different phases of a learning session with MetaTutor. We further explored how the emotion profiles and variable-centered patterns identified through these approaches relate to learning outcomes (i.e., through a latent growth linear mixed effect model), and to self-reported habitual emotion regulations strategies.

With the person-centered approach we identified three distinct emotion profiles that reflected different emotional experiences during learning with MetaTutor. In line with our hypotheses and previous research, these profiles included a positive, negative, and neutral (referred to as low intensity in other studies; Robinson et al., 2017) emotion profile. However, it is important to note that the negative profile was not characterized by high levels of negative emotion intensities. It rather represented a group of students that had higher levels of negative emotions than the students belonging to the other profiles. An exception to this pattern was boredom, as the neutral profile showed comparable levels of boredom. This is in line with findings of previous studies emphasizing the distinct role of boredom during learning (Goetz et al., 2014). These findings were further supported through the variable-centered emotion patterns we identified in subsequent steps. Across six points of time throughout the learning session negative and positive emotions remained separate variable-centered patterns from boredom and neutral. This indicates that the separating features of our emotion profiles are related to a stable cluster structure of emotions. Moreover, our results indicated that the most profound difference in emotional experience between emotion profiles were found for the negative emotions ($\eta^2 = 0.48$ for the negative emotion cluster scores as compared to $\eta^2 = 0.09$ for other emotion cluster scores). In our profile solutions negative emotions were associated with one another regardless of their level of arousal. Interestingly, surprise was associated with the negative profile and negative emotions cluster. This finding corresponds with findings of a previous study that found a significant negative relation between surprise and the accuracy of metacognitive judgements indicating a potential negative impact on learning (Taub et al., in press). However, the lack of differentiation of levels of arousal is likely caused by the imbalanced nature of arousal and valence in emotions measured in the present study (Robinson et al., 2017). Particularly, positive deactivating emotions were underrepresented in the EV. Nonetheless, across two different approaches we identified a theoretically supported and meaningful structure of emotions that centered around three levels of valence—i.e., positive, neutral, and negative.

The most striking feature across all profile solutions was the stability of the negative profile. More specifically, 26 of the 39 (67%) students in the negative profile were always assigned to the same profile regardless of the number of other profiles.[12] This indicates that the group of students with higher levels of negative emotions is most distinct from all other students (in regard to emotional experience). More importantly, comparisons of the learning outcomes for the profiles revealed that the

[12] This pattern was even stronger in the hierarchical cluster analyses as over 90% of students in that profile were consistently assigned to the same profile regardless of the number of other profiles.

negative profile performed significantly worse than at least one other profile at post-test in most profile solutions. In the three-profile solution presented in this paper, the negative profile was significantly outperformed by the neutral profile. This finding is well in line with previous studies using person-centered approaches, as multiple studies found that students with negative emotion profiles tend to learn less than those with neutral or positive profiles (Ganotice Jr et al., 2016; Jarrell et al., 2017; Robinson et al., 2017; see Table 4.2). As opposed to variable-centered approaches that showed positive and negative effects of negative and positive emotions depending on the circumstances, person-centered approaches consistently found detrimental effects of negative emotions for learning. While under certain circumstances single negative (resolved) emotions can potentially benefit learning strategies and outcomes (e.g., D'Mello & Graesser, 2014; Taub et al., in press), our data provided no support for beneficial effects of experiencing multiple negative emotions (e.g., students that belong to a negative emotion profile). It is important to note that while mixed effects of positive and negative emotions depending on the circumstances have been found in multiple studies, most studies indicate that positive emotions are typically beneficial and negative emotions are detrimental for learning (Boekaerts & Pekrun, 2015). Our results supported this general trend for negative emotions.

In addition to the question of which profiles do significantly differ in learning, we also investigated if and how variable-centered emotion patterns would predict learning. We found that positive emotions before the actual learning activity (EVs 1 and 2, see 5.4.1 Emotion items) can predict learning outcomes beyond the explanatory effect of prior knowledge. During self-regulated learning with MetaTutor only negative emotions were significant predictors of learning, but not consistently (significant for EV3 and EV6, only marginally significant for EV4 and EV5). These findings indicate that predictive value of variable-centered emotion patterns for learning fluctuates over time and that negative emotions seem to play a predominant role during the learning activity. Furthermore, these finding reflect central approaches related to learning in digital learning environments – products and processes (Garcia-Martin & Garcia-Sanchez, 2018). More specifically, the profile analysis conducted in this study is primarily product focused as we first investigated differences in learning outcomes (i.e., product data) between emotion profiles. With subsequent analyses, we investigated the process nature of emotions by assessing how emotions form patterns over time and how linear developments in these patterns are related to learning.

We faced several challenges and identified limitations when applying the two clustering approaches to the present data. Our sampling approach was defined relative to the start and end of the session. In particular, we selected the first two EVs and the last two in the learning session. Of these questionnaires, only the first in the learning phase (EV3) and the very last before the posttest (EV6) were administered identically for all participants. The EVs in between these were identical relative to the start and end of the learning session, but slightly different in regard to learning time depending on the total number of EVs the participant completed (e.g., for participants with six EVs all questionnaires were in an actual sequence, while for participants with eight EVs the new sequence included the first four EVs and the last two EVs, leaving two EVs out and creating a spline which might not completely reflect the initial temporal trajectory). However, both profile analyses across all time points and the emotion clusters revealed that the selected clusters represented a stable, comparable selection of measures over time.

As a potential explanation for differences between emotion profiles we compared them in regard to emotion regulation and found significant differences in cognitive reappraisal, but not for expressive suppression between profiles. More specifically, the negative profile reported significantly lower habitual use of cognitive reappraisal than the positive profile, but not compared to the neutral profile. To back up these findings we compared the profiles in regard to variable-centered emotion pattern scores and their linear temporal trajectories. We found that emotion profiles did not only differ in averaged emotion pattern scores for all identified emotion patterns but also exhibited significantly different linear growth for negative emotions, positive emotions and boredom (see Figure 4.4). The most distinct differences lied in the negative emotion pattern as the negative profile displayed a linear increase in negative emotion pattern scores while the scores decreased/stagnated in the other profiles. This illustrates that the negative profile not only starts with higher values of negative emotions, but that this difference got larger over time. Taken together with our finding that the negative emotions cluster negatively predicted learning throughout the learning phase, this indicates that the issues of the negative emotion profile seem to arise over time and are linked to emotion regulation.

A potential explanation for the suboptimal performance of the negative emotion profile is the potential load on working memory imposed by negative emotions and emotion regulation (Curci, Lanciano, Soleti, & Rimé, 2013). While positive emotions cannot enhance working memory beyond its natural capacity, multiple negative emotions may block valuable resources that are particularly required for mastering complex topics and completing challenging learning tasks. This phenomenon might be even more important in digital learning environments as they impose significant challenges to learners (e.g., for navigation through non-linear hyperlinked environments, coordinating multiple goals, integrate agent feedback, use sophisticated learning strategies; Opfermann et al., 2013). Future studies aiming to explain why negative emotions pose a detrimental effect on learning are needed, including cognitive load and its relation to working memory (Anmarkrud, Andresen & Bråten, 2019; Seufert, 2018).

Another limitation of the present study (and person-centered approaches in general) is the decontextualized nature of emotion measures used. Theories on affective dynamics stretch the importance of specific events or impasses that elicit emotions, however the events preceding the measurement of emotions have not been considered yet. Specifically, given our data we cannot disentangle whether students learned less because they experienced negative emotions or if they experienced negative emotions because they were having difficulties during the learning process. Identifying if the elevated levels of negative emotions in negative profiles is related to characteristics of the learning task or the learning environment is crucial for both the understanding of the profiles and the development of adaptive systems that can support students and circumvent negative effects of negative emotions on learning though scaffolds. For instance, in our study we cannot rule out that the increase in negative emotion, especially in the negative emotions profile, was related participants being prompted to fill out self-reports to indicate their emotions repeatedly during the learning activity. Likewise, the precedents of emotional reactions during learning should be incorporated in future studies (e.g., by assessing which emotions specific prompts of pedagogical agents elicit). Taub et al. (in press) have shown that facially expressed emotions are associated with the accuracy of learning strategies. Identifying arising negative emotions and the learning processes they directly affect can bridge the gap between emotions and (meta)-cognitive processes. This goes hand in hand with another shortcoming of this line of inquiry – the sole reliance on self-reports to measure emotions. Models and research on emotions clearly state that emotions are multi-faceted processes and limiting our scope to the appraisal component (Scherer & Moors, 2019) is a significant limitation. Building multi-channel, multi-modal emotion profiles through the use of additional data channels can benefit person-centered research by refining profiles and by providing additional explanations how the profiles develop over time (e.g., through peaks in EDA). Lastly, personal predispositions (e.g., personality – narcissism as a predisposition for negative emotionality) is a general cause for differences in emotional experience and emotion regulation, and its effect on learning strategies could be very beneficial to deepen the understanding of emotions in self-regulated learning processes.

4.7 Conclusion

In conclusion, the results of our study highlight the importance of negative emotions during self-regulated learning with digital learning environments during complex learning. The present study adds to research in multiple ways. Methodologically, we have showcased how a person-centered and a novel variable-centered approach complement each other. Particularly identifying variable-centered emotion patterns in addition to emotion profiles enabled us to analyze temporal dynamics of multiple emotions simultaneously. A negative relation between negative emotions and learning outcomes was found with both approaches. This underlines the robustness of this finding and further shows that person-centered and variable-centered approaches can supplement each other. Moreover, clustering approaches offer the possibility to further connect findings from studies using different measures more easily (e.g., achievement emotions vs. learning-centered emotions). Through the combination of person-centered and variable-centered approaches, we have found that both the students with the highest levels of negative emotions overall and higher levels of negative emotions across all students showed a significant negative relation to learning. Furthermore, we have found that these detrimental effects are linked to lower (self-reported) emotion regulation. This indicates the need to identify when elevated levels of negative emotions arise, particularly for students who experience a multitude of negative emotions, for practitioners and researchers to intervene in a timely fashion before the detrimental effects of negative emotions settle in. Specifically, fostering students' emotion regulation as part of self-regulated learning activities with digital learning environments is a promising prospect to improve students' emotional experience and learning subsequently. Therefore, the design, development, and implementation of digital learning environments as well as educational interventions should incorporate emotions and emotion regulation as parts of (self-regulated) learning activities to maximize positive effects on students' learning.

References

Anmarkrud, Ø., Andresen, A., & Bråten, I. (2019). Cognitive Load and Working Memory in Multimedia Learning: Conceptual and Measurement Issues. *Educational Psychologist*, 1-23.

Arguel, A., Lockyer, L., Kennedy, G., Lodge, J. M., & Pachman, M. (2019). Seeking optimal confusion: a review on epistemic emotion management in interactive digital learning environments. *Interactive Learning Environments, 27*(2), 200-210.

Asendorpf, J. B., Borkenau, P., Ostendorf, F., & Van Aken, M. A. G. (2001). Carving personality description at its joints: Confirmation of three replicable personality prototypes for both children and adults. *European Journal of Personality, 15*(3), 169-198. doi:DOI 10.1002/per.408.abs

Azevedo, R., Harley, J., Trevors, G., Duffy, M., Feyzi-Behnagh, R., Bouchet, F., & Landis, R. (2013). Using trace data to examine the complex roles of cognitive, metacognitive, and emotional self-regulatory processes during learning with multi-agent systems. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (Vol. 28, pp. 427-449). New York, NY: Springer.

Azevedo, R., Martin, S. A., Taub, M., Mudrick, N. V., Millar, G. C., & Grafsgaard, J. F. (2016). Are Pedagogical Agents' External Regulation Effective in Fostering Learning with Intelligent Tutoring Systems? In A. Micarelli, J. Stamper, & K. Panourgia (Eds.), *Intelligent Tutoring Systems. ITS 2016* (pp. 197-207). Zagreb, Croatia: Springer.

Azevedo, R., Mudrick, N. V., Taub, M., & Bradbury, A. (2019). Self-regulation in computer-assisted learning systems. In J. Dunlosky & K. Rawson (Eds.), *Handbook of Cognition and Education* (pp. 587-618). Cambridge, MA: Cambridge University Press.

Azevedo, R., Mudrick, N. V., Taub, M., & Wortha, F. (2017). Coupling between metacognition and emotions during STEM learning with advanced learning technologies: A critical analysis, implications for future research, and design of learning systems. In T. Michalsky & C. Schechter (Eds.), *Self-regulated learning: Conceptualization, contribution, and empirically based models for teaching and learning* (pp. 1-18). New York, NY: Teachers College Press.

Azevedo, R., Taub, M., & Mudrick, N. V. (2018). Using multi-channel trace data to infer and foster self-regulated learning between humans and advanced learning technologies. In D. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 254-270). New York, NY: Routledge.

Baker, R. S., D'Mello, S. K., Rodrigo, M. M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies, 68*(4), 223-241.

Barker, E. T., Howard, A. L., Galambos, N. L., & Wrosch, C. (2016). Tracking affect and academic success across university: Happy students benefit from bouts of negative mood. *Developmental Psychology, 52*(12), 2022-2030.

Baumeister, R. F., Alquist, J. L., & Vohs, K. D. (2015). Illusions of learning: Irrelevant emotions inflate judgments of learning. *Journal of Behavioral Decision Making, 28*(2), 149-158.

Behrendt, S. (2014). lm.beta: Add standardized regression coefficients to lm-objects. Retrieved from https://CRAN.R-project.org/package=lm.beta

Ben-Eliyahu, A., & Linnenbrink-Garcia, L. (2013). Extending self-regulated learning to include self-regulated emotion strategies. *Motivation and Emotion, 37*(3), 558-573.

Boekaerts, M., & Pekrun, R. (2015). Emotions and emotion regulation in academic settings. In L. Corno & E. M. Anderman (Eds.), *Handbook of Educational Psychology* (pp. 76-90). New York, NY: Routledge.

Breckenridge, J. N. (2000). Validating cluster analysis: Consistent replication and symmetry. *Multivariate Behavioral Research, 35*(2), 261-285.

Charrad, M., Ghazzali, N., Boiteau, V., Niknafs, A., & Charrad, M. M. (2014). Package 'nbclust'. *Journal of Statistical Software, 61*, 1-36.

Craig, S., Graesser, A., Sullins, J., & Gholson, B. (2004). Affect and learning: an exploratory look into the role of affect in learning with AutoTutor. *Journal of Educational Media, 29*(3), 241-250.

Curci, A., Lanciano, T., Soleti, E., & Rimé, B. (2013). Negative emotional experiences arouse rumination and affect working memory capacity. *Emotion, 13*(5), 867-880. doi:10.1037/a0032492

D'Mello, S. (2013). A selective meta-analysis on the relative incidence of discrete affective states during learning with technology. *Journal of Educational Psychology, 105*(4), 1082.

D'Mello, S., & Graesser, A. (2014). Confusion and its dynamics during device comprehension with breakdown scenarios. *Acta Psychologica, 151*, 106-116.

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction, 22*(2), 145-157.

D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction, 29*, 153-170.

Ekman, P. (1992). An argument for basic emotions. *Cognition & emotion, 6*(3-4), 169-200.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17*(2), 124-129.

Eshghi, A., Haughton, D., Legrand, P., Skaletsky, M., & Woolford, S. (2011). Identifying groups: A comparison of methodologies. *Journal of Data Science, 9*(2), 271-292.

Fernando, J. W., Kashima, Y., & Laham, S. M. (2014). Multiple emotions: A person-centered approach to the relationship between intergroup emotion and action orientation. *Emotion, 14*(4), 722.

Fortunato, V. J., & Goldblatt, A. M. (2006). An examination of goal orientation profiles using cluster analysis and their relationships with dispositional characteristics and motivational response patterns. *Journal of Applied Social Psychology, 36*(9), 2150-2183.

Garcia-Martin, J., & Garcia-Sanchez, J. N. (2018). The instructional effectiveness of two virtual approaches: processes and product. *Revista de Psicodidáctica (English ed.)*, *23*(2), 117-127.

Ganotice Jr, F. A., Datu, J. A. D., & King, R. B. (2016). Which emotional profiles exhibit the best learning outcomes? A person-centered analysis of students' academic emotions. *School Psychology International, 37*(5), 498-518.

Gegenfurtner, A., Fryer, L., Järvelä, S., Harackiewicz, J., & Narciss, S. (2019). Affective learning in digital education. *Frontiers in Education*.

George, D., & Mallery, P. (1999). *SPSS® for Windows® step by step: A simple guide and reference*: Needham Heights, MA: Allyn & Bacon.

Goetz, T., Frenzel, A. C., Hall, N. C., Nett, U. E., Pekrun, R., & Lipnevich, A. A. (2014). Types of boredom: An experience sampling approach. *Motivation and Emotion, 38*(3), 401-419.

Graesser, A. C., Rus, V., D'Mello, S., & Jackson, G. (2008). AutoTutor: Learning through natural language dialogue that adapts to the cognitive and affective states of the learner. In D. H. Robinson & G. Schraw (Eds.), *Current perspectives on cognition, learning and instruction: Recent innovations in educational technology that facilitate student learning* (pp. 95-125). Charlotte, NC: Information Age Publishing.

Grafsgaard, J., Wiggins, J., Vail, A., Boyer, K., Wiebe, K., & Lester, J. (2014). The additive value of multimodal features for predicting engagement, frustration, and learning during tutoring. In A. A. Salah, J. Cohn & B. Schuller (Eds.) *Proceedings of the Sixteenth ACM International Conference on Multimodal Interaction* (pp. 42–49) New York, NY: ACM.

Gross, J. J. (2015). Emotion regulation: Current status and future prospects. *Psychological Inquiry, 26*(1), 1-26.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*(2), 348-362.

Grubbs, F. E. (1969). Procedures for detecting outlying observations in samples. *Technometrics, 11*(1), 1-21.

Harley, J. M., Pekrun, R., Taxer, J. L., & Gross, J. J. (2019). Emotion regulation in achievement situations: An Integrated Model. *Educational Psychologist, 54*(2), 106-126. doi:10.1080/00461520.2019.1587297

Harley, J. M., Taub, M., Azevedo, R., & Bouchet, F. (2017). Let's set up some subgoals: Understanding human-pedagogical agent collaborations and their implications for learning and prompt and feedback compliance. *IEEE Transactions on Learning Technologies, 11*(1), 54-66.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical journal, 50*(3), 346-363.

Jarrell, A., Harley, J. M., Lajoie, S., & Naismith, L. (2017). Success, failure and emotions: examining the relationship between performance feedback and emotions in diagnostic reasoning. *Educational Technology Research and Development, 65*(5), 1263-1284.

Jarrell, A., Harley, J. M., & Lajoie, S. P. (2016). The link between achievement emotions, appraisals, and task performance: pedagogical considerations for emotions in CBLEs. *Journal of Computers in Education, 3*(3), 289-307.

Kluger, Y., Basri, R., Chang, J. T., & Gerstein, M. (2003). Spectral biclustering of microarray data: coclustering genes and conditions. *Genome Research, 13*(4), 703-716.

Komsta, L. (2011). outliers: Tests for outliers. Retrieved from https://CRAN.R-project.org/package=outliers

Lazarus, R. S. (2006). Emotions and interpersonal relationships: Toward a person-centered conceptualization of emotions and coping. *Journal of Personality, 74*(1), 9-46.

Lewis, M., Haviland-Jones, J. M., & Barrett, L. F. (2008). *Handbook of emotions, 3rd ed*. New York, NY: The Guilford Press.

Makowski, (2018). The psycho package: An efficient and publishing-oriented workflow for psychological science. Journal of Open Source Software, 3(22), 470. https://doi.org/10.21105/joss.00470

Moors, A., Ellsworth, P. C., Scherer, K. R., & Frijda, N. H. (2013). Appraisal theories of emotion: State of the art and future development. *Emotion Review, 5*(2), 119-124.

Muis, K. R., Pekrun, R., Sinatra, G. M., Azevedo, R., Trevors, G., Meier, E., & Heddy, B. C. (2015). The curious case of climate change: Testing a theoretical model of epistemic beliefs, epistemic emotions, and complex learning. *Learning and Instruction, 39*, 168-183.

Opfermann, M., Scheiter, K., Gerjets, P., & Schmeck, A. (2013). Hypermedia and self-regulation: An Interplay in both directions. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 129-141). New York, NY: Springer New York.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825-2830.

Pekrun, R. (2006). The Control-Value Theory of Achievement Emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315-341. doi:10.1007/s10648-006-9029-9

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

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child development, 88*(5), 1653-1670.

Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 259-282). Boston, MA: Springer.

Pekrun, R., & Linnenbrink-Garcia, L. (2014). *International handbook of emotions in education*. New York, NY: Routledge/Taylor & Francis Group.

Pekrun, R., Vogl, E., Muis, K. R., & Sinatra, G. M. (2017). Measuring emotions during epistemic activities: the Epistemically-Related Emotion Scales. *Cognition and Emotion, 31*(6), 1268-1276.

Plass, J. L., Heidig, S., Hayward, E. O., Homer, B. D., & Um, E. (2014). Emotional design in multimedia learning: Effects of shape and color on affect and learning. *Learning and Instruction, 29*, 128-140.

R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing: Vienna, Austria. URL http://www.R-project.org/

Revelle, W. R. (2017). psych: Procedures for personality and psychological research (Version 1.8.12). Retrieved from https://CRAN.R-project.org/package=psych

Riemer, V., & Schrader, C. (2019). Mental model development in multimedia learning: Interrelated effects of emotions and self-monitoring. *Frontiers in Psychology, 10*(899). doi:10.3389/fpsyg.2019.00899

Robinson, K. A., Ranellucci, J., Lee, Y.-k., Wormington, S. V., Roseth, C. J., & Linnenbrink-Garcia, L. (2017). Affective profiles and academic success in a college science course. *Contemporary Educational Psychology, 51*, 209-221.

Sabourin, J. L., & Lester, J. C. (2014). Affect and engagement in game-based learning environments. *IEEE Transactions on Affective Computing, 5*(1), 45-56.

Scherer, K., & Moors, (2019). The emotion process: Event appraisal and component differentiation. *Annual Review of Psychology, 70*, 719-745.

Seufert, T. (2018). The interplay between self-regulation in learning and cognitive load. *Educational Research Review*, *24*, 116-129.

Sinclair, J., Jang, E. E., Azevedo, R., Lau, C., Taub, M & Mudrick, N. (2018). Changes in emotion and their relationship with learning gains in the context of MetaTutor. In R. Nkambou, R. Azevedo, & J. Vassileva (Eds.), *Lecture notes in computer science: Intelligent tutoring systems 2018* (pp. 202-211). Cham: Springer.

Spann, C. A., Shute, V. J., Rahimi, S., & D'Mello, S. K. (2019). The productive role of cognitive reappraisal in regulating affect during game-based learning. *Computers in Human Behavior*. *100*, 358-369.

SpSS, I. (2012). SPSS version 21.0. *IBM SPSS, Chicago, Illinois, USA*.

Taub, M., Azevedo, R., Bouchet, F., & Khosravifar, B. (2014). Can the use of cognitive and metacognitive self-regulated learning strategies be predicted by learners' levels of prior knowledge in hypermedia-learning environments?. *Computers in Human Behavior*, *39*, 356-367.

Taub, M., Azevedo, R., Bradbury, A. E., Millar, G. C., & Lester, J. (2018). Using sequence mining to reveal the efficiency in scientific reasoning during STEM learning with a game-based learning environment. *Learning and Instruction, 54*, 93-103.

Taub, M., Azevedo, R., Rajendran, R., Cloude, E. B., Biswas, G., & Price, M. J. (in press/online first 2019). How are students' emotions related to the accuracy of cognitive and metacognitive processes during learning with an intelligent tutoring system? *Learning and Instruction*.

Tyng, C. M., Amin, H. U., Saad, M. N. M., & Malik, A. S. (2017). The influences of emotion on learning and memory. *Frontiers in Psychology, 8*(1454). doi:10.3389/fpsyg.2017.01454

Van Rossum, G., & Drake, F. L. (2011). *Python Language reference manual*. Bristol, UK: Network Theory Ltd.

Vansteenkiste, M., Sierens, E., Soenens, B., Luyckx, K., & Lens, W. (2009). Motivational profiles from a self-determination perspective: The quality of motivation matters. *Journal of Educational Psychology, 101*(3), 671-688.

Venables, W., & Ripley, B. (2002). *Statistics complements to modern applied statistics with S (4th edition)*. New York, NY: Springer.

Zheng, L. (2016). The effectiveness of self-regulated learning scaffolds on academic performance in computer-based learning environments: A meta-analysis. *Asia Pacific Education Review, 17*(2), 187-202.

Table 4.1

Overview of person-centered studies on emotions during learning

Study	Sample	Clustering variables	Identified clusters (method)	Main findings
Jarrell et al., 2016	Medical students (N = 26)	Enjoyment, pride, hope, shame, and anger	3 (k-means clustering) • Positive • Negative • Low	No significant differences in performance between profiles
Jarrell et al., 2017	Medical / Dentistry students (N = 30)	Enjoyment, pride, hope, shame, and anger	3 (k-means clustering) • Positive • Negative • Low	Negative profile is significantly outperformed by at least one other cluster
Ganotice Jr et al., 2016	Secondary school students (N_1 = 1,147; N_2 = 341)	Enjoyment, hope, pride, anger, anxiety, shame, hopelessness, and boredom	4 (hierarchical + k-means clustering) • high positive and high shame • moderate positive and negative • high negative • high positive emotion	High positive emotions cluster showed best academic outcomes High negative emotions cluster showed worst academic outcomes
Robinson et al., 2017	Undergraduate students (N = 278)	Affect: positive/negative x activated/deactivated	4 (hierarchical + k-means clustering) • Positive • Deactivated • Negative • Moderate negative	Deactivated profile showed higher academic achievement than both negative profiles
Sinclair et al., 2018	Undergraduate students (N = 190)	Enjoyment, curiosity, pride, boredom, and frustration	3 (Latent profile analysis) • Positive • Negative (bored/frustrated) • Moderate	Students in the negative profile were least likely to change to another profile Learning gains are associated with transitions between profiles

Table 4.3

Means and standard deviations for emotion items, emotion regulation, and learning measures by profile solutions

Profile solution	3			4				5				
Profile	1	2	3	1	2	3	4	1	2	3	4	5
n	75	62	39	27	67	50	32	27	60	27	32	30
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Enjoyment	2.41 (0.67)	3.73 (0.62)	2.5 (0.66)	1.66 (0.42)	2.84 (0.53)	3.84 (0.59)	2.6 (0.59)	1.66 (0.42)	2.8 (0.51)	3.55 (0.63)	2.6 (0.59)	3.94 (0.62)
Hope	2.8 (0.71)	4.05 (0.55)	2.83 (0.64)	2.29 (0.66)	3.09 (0.61)	4.19 (0.51)	2.92 (0.58)	2.29 (0.66)	3.10 (0.6)	3.69 (0.63)	2.92 (0.58)	4.37 (0.49)
Pride	2.21 (0.77)	3.43 (0.78)	2.55 (0.65)	1.78 (0.55)	2.50 (0.74)	3.53 (0.81)	2.67 (0.63)	1.78 (0.55)	2.5 (0.71)	3.03 (0.96)	2.67 (0.63)	3.75 (0.67)
Frustration	2.04 (0.84)	1.63 (0.56)	2.91 (0.61)	2.72 (0.87)	1.76 (0.68)	1.67 (0.56)	2.92 (0.57)	2.72 (0.87)	1.77 (0.68)	1.93 (0.56)	2.92 (0.57)	1.43 (0.47)
Anxiety	1.91 (0.81)	2.11 (1.00)	3.04 (0.76)	2.12 (0.81)	1.88 (0.81)	2.16 (1.01)	3.19 (0.75)	2.12 (0.81)	1.76 (0.65)	3.00 (0.84)	3.19 (0.75)	1.57 (0.74)
Shame	1.23 (0.34)	1.19 (0.3)	2.26 (0.56)	1.29 (0.34)	1.22 (0.35)	1.2 (0.32)	2.41 (0.47)	1.29 (0.34)	1.19 (0.28)	1.48 (0.45)	2.41 (0.47)	1.02 (0.05)
Hopelessness	1.23 (0.33)	1.15 (0.26)	2.38 (0.55)	1.31 (0.38)	1.24 (0.34)	1.14 (0.25)	2.52 (0.47)	1.31 (0.38)	1.23 (0.34)	1.30 (0.34)	2.52 (0.47)	1.04 (0.09)
Boredom	3.19 (0.93)	1.99 (0.68)	3.76 (0.67)	4.24 (0.56)	2.73 (0.73)	1.91 (0.67)	3.66 (0.63)	4.24 (0.56)	2.83 (0.68)	1.88 (0.63)	3.66 (0.63)	1.93 (0.68)
Surprise	1.46 (0.48)	1.93 (0.88)	2.2 (0.68)	1.23 (0.33)	1.57 (0.49)	2.04 (0.91)	2.33 (0.65)	1.23 (0.33)	1.56 (0.49)	2.56 (0.8)	2.33 (0.65)	1.47 (0.59)
Contempt	1.53 (0.74)	1.92 (1.05)	2.74 (0.66)	2.01 (0.93)	1.60 (0.87)	1.88 (1.08)	2.65 (0.55)	2.01 (0.93)	1.63 (0.9)	1.88 (0.89)	2.65 (0.55)	1.75 (1.16)
Confusion	1.38 (0.43)	1.49 (0.53)	2.41 (0.56)	1.55 (0.65)	1.40 (0.43)	1.51 (0.55)	2.45 (0.55)	1.55 (0.65)	1.40 (0.42)	1.86 (0.54)	2.45 (0.55)	1.18 (0.3)

Table 4.3 (continued).

Profile solution	3			4				5				
Profile	1	2	3	1	2	3	4	1	2	3	4	5
Curiosity	2.50	3.75	2.94	1.93	2.82	3.89	3.10	1.93	2.82	4.02	3.10	3.54
	(0.66)	(0.81)	(0.76)	(0.59)	(0.62)	(0.74)	(0.64)	(0.59)	(0.59)	(0.58)	(0.64)	(0.95)
Sadness	1.19	1.17	2.09	1.28	1.16	1.19	2.23	1.28	1.16	1.29	2.23	1.10
	(0.37)	(0.32)	(0.61)	(0.42)	(0.34)	(0.35)	(0.55)	(0.42)	(0.34)	(0.43)	(0.55)	(0.22)
Neutral	3.70	3.15	3.19	3.30	3.79	3.03	3.20	3.30	3.90	2.79	3.20	3.20
	(0.71)	(0.83)	(0.75)	(0.86)	(0.61)	(0.84)	(0.7)	(0.86)	(0.53)	(0.77)	(0.7)	(0.8)
Reappraisal	4.95	5.30	4.62	5.03	5.00	5.36	4.41	5.03	5.03	5.16	4.41	5.38
	(1.09)	(1.07)	(1.27)	(1.26)	(0.92)	(1.15)	(1.30)	(1.26)	(0.88)	(1.22)	(1.30)	(1.13)
Suppression	3.97	3.97	3.94	3.78	3.97	4.03	4.01	3.78	4.03	4.09	4.01	3.83
	(0.97)	(1.13)	(1.17)	(0.98)	(1.04)	(1.11)	(1.16)	(0.98)	(1.04)	(1.05)	(1.16)	(1.15)
Pre ratio	0.59	0.58	0.55	0.53	0.60	0.59	0.55	0.53	0.60	0.59	0.55	0.59
	(0.13)	(0.14)	(0.17)	(0.11)	(0.13)	(0.13)	(0.19)	(0.11)	(0.13)	(0.12)	(0.19)	(0.14)
Post ratio	0.71	0.70	0.63	0.67	0.72	0.7	0.62	0.67	0.72	0.71	0.62	0.70
	(0.12)	(0.13)	(0.16)	(0.10)	(0.13)	(0.13)	(0.17)	(0.10)	(0.13)	(0.11)	(0.17)	(0.14)

Note. Reappraisal = cognitive reappraisal subscale of the emotion regulation questionnaire, suppression = expressive suppression subscale of the emotion regulation questionnaire

Table 4.4

Maintained variance and loadings for emotion patterns

Pattern	Variable	Time point						
		Overall	EV T1	EV_T2	EV_T3	EV_T4	EV_T5	EV_T6
Negative	σ²	0.40	0.38	0.35	0.40	0.43	0.42	0.50
	frustration	0.85	0.52	0.56	0.89	1.02	0.94	0.87
	anxiety	0.83	0.95	0.74	0.82	0.86	0.84	0.95
	shame	0.60	0.50	0.59	0.61	0.68	0.50	0.67
	hopelessness	0.63	0.44	0.40	0.71	0.72	0.71	0.70
	surprise	0.43	0.45	0.52	0.53	0.32	0.36	0.45
	confusion	0.64	0.39	0.47	0.68	0.67	0.70	0.80
	sadness	0.49	0.36	0.49	0.49	0.52	0.54	0.50
	contempt*	0.52	0.58	0.54	0.65	0.49	0.65	
Positive	σ²	0.65	0.55	0.68	0.65	0.66	0.63	0.66
	enjoyment	0.94	0.73	0.85	0.94	0.92	1.06	1.02
	hope	0.97	0.78	0.90	0.92	0.95	1.02	1.02
	pride	0.86	0.77	0.89	0.89	0.88	0.87	0.98
	curiosity	0.99	0.57	0.88	0.95	0.94	0.87	0.89
Neutral	σ²	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	neutral	1.16	1.08	1.15	1.06	1.12	1.15	1.15
Boredom	σ²	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	boredom	1.34	1.18	1.22	1.36	1.41	1.39	1.27
	contempt*							0.61

Note. σ²: Maintained variance. EV: Emotion values questionnaire. Absolute loading values were used if all loadings on the same main component were negative. *Contempt at EV T6 is the only deviation from the stable structure of emotions. It was associated with the boredom pattern instead of the negative pattern.

5 General Discussion

5.1 Discussion of general findings

The ability to effectively self-regulate learning processes is essential to succeed in modern educational settings. Recent developments, such as the heavily increased use of information technology in autonomous learning activities (e.g., online classes and courses) have further increased the importance of SR to achieve desirable learning outcomes. Yet, based on a variety of diverse perspectives though which SR has been researched it remains unclear which specific processes SR entails. This dissertation aimed to address this issue by providing an integrative framework that expands upon the core processes of SRL (i.e., learning activities) with the differentiation and inclusion of driving forces, personal dispositions, and limited resources. The three studies in this dissertation provided empirical evidence for the value of this framework when investigating SRL from a broader perspective. Particularly, the predictive value of the different areas of this framework was tested for different outcomes across different levels of granularity. In support of the proposed framework, the first study showed that all areas of the model provide important contributions to the prediction of grades and laboratory learning outcomes across five content domains. Further, this study showed that predictions in both contexts (i.e., school and laboratory) showed similarities in patterns, but distinct differences in the most predictive variables. The second study underlined the importance of all components of the framework in an art-learning task and further demonstrated how the learning medium (i.e., PC or tablet) affected the predictive value of self-regulatory constructs. Finally, the complex, temporal dynamics of emotions as a driving force during learning in a hypermedia-based tutoring system were linked to learning outcomes and support for personal dispositions as an underlying cause was revealed (Study III). In the next section the value of jointly investigating all areas of the proposed framework (chapter 5.1) as well as the relative importance of each area will be discussed (chapter 5.2). Subsequently, strength and limitations of this dissertation will be outlined and implications for future research and practice will be derived (chapter 5.3).

5.1.1 A bigger picture of self-regulation in education

This dissertation builds upon the common notion of SRL as a complex, multi-faceted process. According to this view SRL is typically described as the regulation of thoughts, behaviors, emotions, and motivation (e.g., Schunk & Greene, 2018a) or cognitive, affective, metacognitive and motivational processes (e.g., Duffy & Azevedo, 2015). Due to this broad conceptualization, investigations on SRL are informed by and directly related to many fields of research in education and beyond. To consolidate widely differing perspectives, I proposed an integrative framework, informed by previous research categorizing predictors of learning (e.g., Richardson et al., 2012). Specifically, self-regulatory constructs were categorized into learning activities, driving forces, personal dispositions, and limited resources based on the mechanism through which they affect learning. Evidence from the individual lines of research clearly showed that learning activities, driving forces, personal dispositions, and limited resources by their own are strong, distinctive determinants of learning and academic achievement. Yet, the question how they jointly predict performance, to my knowledge, has not been comprehensively researched.

The overall finding of this dissertation is that predictors from all areas of the proposed framework are required to optimally predict learning in different settings and different levels of granularity support the underlying structure of the proposed framework. Particularly, the consistent finding that patterns of learning activities, driving forces, personal dispositions, and limited resources predicted learning outcomes in school as well as across (Study I) and within specific laboratory learning tasks (Study II), demonstrated that predictors from different areas of SR in education add value to one another, rather than a single research tradition dominating predictions. Further, Study III investigated the temporal unfolding of driving forces and their relation to learning and personal dispositions. Results showed that the effect of driving forces on learning might be based in stable personal dispositions, by demonstrating that participants with negative emotionality during learning are also defined by a low habitual use of emotion regulation strategies. Further, for this dissertation, we extended these analyses by comparing the identified emotion profiles with regards to their personal dispositions (Donnellan, Oswald, Baird, & Lucas, 2006). Specifically, analyses showed that negative emotion profiles were characterized by significantly higher levels of neuroticism than other profiles but showed no differences in other personality traits. In sum, throughout all studies results showed that learning activities, driving forces, personal dispositions, and limited resources provide complementary value for the prediction of learning across different levels of granularity and contexts. This finding builds on initial studies that aimed to bridge two of the four components proposed in this framework (e.g., Bidjerano & Dai, 2007; de Bruin & van Merriënboer, 2017; Follmer & Sperling, 2016). However, this dissertation presents the first empirical approach to integrate research on SR in education. Therefore, the studies in this dissertation pioneered the impact of constructs from different research traditions on learning when an extensive set of explanatory mechanisms is considered. No systematic difference in the predictive value based on the type of measure was revealed, as opposed to attempts to integrate SR constructs in clinical settings, which found that cognitive task data showed weaker relations to life outcomes than self-report data (Eisenberg et al., 2019). For instance, robust predictions of grades and laboratory learning task performance (Study I) included self-report measures (e.g., expectancy value) as well as different measures of cognitive resources obtained from cognitive tasks (e.g., d' scores from working memory tasks). Despite methodological concerns that the type of measurement determines its predictive value for specific outcomes (Dang et al., 2020), measures representing different theoretical and methodological approaches directly related to learning outcomes in school and laboratory tasks, demonstrating the importance of all components of the proposed framework across contexts.

Taken together, these results provide first evidence that, when investigated jointly, all proposed areas of SR significantly relate to learning outcomes. This supports the assumption that SR in education compromises learning activities, driving forces, personal dispositions, and limited resources as distinct but interrelated components of SR in education. In the following sections the relative merit of each area of the proposed framework will be discussed in light of their relation to learning in different tasks and environments.

5.1.2 The relative importance of self-regulatory constructs across contexts

Across the three studies of this dissertation self-regulatory constructs stemming from four interrelated areas of research on SR in education have been investigated.

Particularly, their predictive value for learning outcomes at different levels of granularity has been extensively tested. While the previous section focused on the merits of jointly investigating learning activities, driving forces, personal dispositions, and limited resources, next the respective findings for each of these research traditions in this dissertation will be discussed.

5.1.2.1 Learning activities

SRL revolves around the deployment, control and adaption of cognitive and metacognitive learning strategies. Accordingly, their predictive value for learning in different contexts was tested in the present dissertation. Findings regarding cognitive learning strategies will be discussed, before results for metacognitive learning strategies are outlined. In the first study, low-performing students in laboratory tasks were characterized by higher levels of self-reported use of rehearsal and memorization strategies than their high-performing counterparts. In other words, underachieving students across multiple laboratory-learning tasks commonly relied on learning strategies that are deemed suboptimal (Bjork, Dunlosky, & Kornell, 2013). This overreliance on ineffective, shallow learning strategies is a common issue of struggling students that SRL interventions and scaffolds aim to address (Azevedo, 2005; Azevedo et al., 2013; Narciss, Proske, & Koerndle, 2007). To efficiently achieve high learning outcomes active and elaborated processing is required (Lockhart & Craik, 1990). The second study was in line with this central assumption by showing that high and low performing students using tablets were distinguished through their use of critical evaluation, an elaborate learning strategy (Boerner, Seeber, Keller, & Beinborn, 2005). Profiting from complex learning environments similar to the one used in this study requires learners to integrate information through elaborative strategies (Kornmann, Kammerer, Zettler, Trautwein, & Gerjets, 2016).

In addition to cognitive learning strategies, learning activities incorporate the use metacognitive strategies and their accuracy. In the studies of this dissertation the most robust relation between metacognition and learning outcomes was found for judgments of learning in a complex, exploratory art learning task (Study II). Further, the use of metacognitive control strategies differed between high and low performing students in school, but the strength of this relation was less pronounced. In light of criticisms regarding the measurement of SRL processes though self-reports (e.g., Winne & Perry, 2000) these findings are in line with expectations. Whereas questionnaires about the use of metacognitive strategies during learning generally asses if and how frequently students deploy them (e.g., Boerner et al., 2005), concrete metacognitive judgments and their accuracy cannot be captured with these measures. Yet, adequate adaptions of learning processes hinge on the in situ use and accuracy of metacognitive judgments (Greene & Azevedo, 2009; Hacker et al., 2009; Metcalfe, 2009). Thus, differences between the implied importance of metacognitive strategies within the framework on SR in education and their predictive value in the studies of this dissertation is likely related to the way they were measures (Jansen et al., 2019). Instead, for performance in school settings, resource management strategies (i.e., effort-related strategies) showed consistent predictive value (Jansen et al., 2019; Theobald, 2021). Together the most predictive learning activities found throughout the studies of this dissertation directly reflect the most (e.g., critical reflection) and the least effective learning strategies (e.g., rehearsal) identified in meta-analytic analyses of SRL in university students (Broadbent & Poon, 2015), further substantiating the robustness of the results in this population.

In conclusion, learning activities in the studies of this dissertation showed that elaborate and effortful learning strategies enabled high learning outcomes, while relying on shallow, surface-level strategies impeded performance. The importance and effectiveness of specific learning strategies varied across learning tasks, domains and other contextual factors (see also, Alexander et al., 2011; Greene et al., 2015). Depending on the learning outcome (i.e., grades vs. laboratory task performance) and context (i.e., learning with PC or tablet) different strategies were linked to performance. More importantly, at least one learning activity was among the strongest predictors of learning when all areas of the proposed framework were considered.

5.1.2.2 Driving forces

SRL processes include a range of driving forces that can catalyze engagement in and maintenance of learning processes, but can further require regulation themselves (e.g., regulation of emotion). In this dissertation three prominent constructs from this research tradition and their relation to learning were tested. Specifically, motivation, interest, as well as emotions and their regulation were investigated as a part of the integrative framework on SR in education. In line with a long tradition of motivational researched based on the expectancy value theory of achievement motivation (Eccles & Wigfield, 2002; Wigfield & Eccles, 2000), motivational variables were robust predictors of learning in school and laboratory tasks. Specifically, two central components form the expectancy value framework, the expectancy x value interaction (Nagengast et al., 2011) and the motivational cost component (Flake, Barron, Hulleman, McCoach, & Welsh, 2015) showed close relations to learning outcomes. In Study I the expectancy x value interaction across the five content domains, was imperative in predictions of grades and consistently related to laboratory task performance, but less impactful in these predictions. This showed, other than the learning strategies discussed in the previous section, a specific driving force determined learning across contexts (i.e., learning in school and laboratory tasks). This underlines the proposed essential role of motivation in SRL process (Boekaerts, 1996; Pintrich, 2000). However, Study II demonstrated that when investigating a specific task in a content domain, the importance of different driving forces can fluctuate based on contextual factors. While interest in art was the most important driving force when learning with a PC, for learners with tablet the motivational cost component (Flake et al., 2015) was consistently related to learning outcomes. Domain-specific interest typically explains differences in academic performance within and across domains (Jansen et al., 2016). Yet, the increased self-regulatory demands shown for the complex art-learning task on tablet was more evident in negative motivational components (i.e., cost). This indicated that, similar to learning activities, driving forces can function as enablers (e.g., interest) or deterrent (e.g., cost) for learning. The later line of argumentation was further underlined in the third study of this dissertation. In a fine-granular analysis of emotions during a learning task the detrimental role of negative emotions for learning was shown, which mirrors common findings and assumptions from research on emotions in education (Pekrun, Goetz, Titz, & Perry, 2002; Pekrun, Lichtenfeld, Marsh, Murayama, & Goetz, 2017). More important in the context of SR in education, was the finding that the negative emotional experience became more pronounced for affected students during learning, which showed close links to suboptimal performance. In line with research on the dynamics of emotions during learning and their regulation (D’Mello & Graesser, 2012; Harley et al., 2019) this underlines the dynamic nature driving forces as processes.

In sum, driving forces showed strong relations to learning across levels of granularity, tasks, and environments. For investigations on a broader level of granularity (i.e., grades and laboratory task performance across multiple domains) dominant driving forces are shared (i.e., expectancy value). In more fine-granular analyses across contexts, driving forces remain an essential component of SR processes, but the specific driving force related to performance varies, similar to learning activities. More importantly, driving forces were powerful predictors of learning across all studies of this dissertation.

5.1.2.3 Personal dispositions

The unfolding of SR processes in learning situations is reliant on the learners' stable personal dispositions. All studies of this dissertation provided empirical support in favor of dispositional of measures of SR. The most consistent relation between personal dispositions and learning outcomes was found for openness. Specifically, this personality facet was among the most frequently selected predictors of performance across five laboratory tasks in Study I and in the art learning task of Study II, particularly for learners using tablets. Despite the predominant focus on conscientiousness and grid as SR dispositions in educational research (e.g., Eilam et al., 2009; Rimfeld et al., 2016) this finding is in line with the literature. Together with conscientiousness, openness has the strongest relation to learning and academic outcomes among the big five personality facets (Gatzka, 2021; Gatzka & Hell, 2018; Vedel & Poropat, 2017). Interestingly, in both studies openness was more closely related to performance in less extreme comparisons. In other words, openness was less important to differentiate the very top and bottom of learners, but consistently included when larger ranges of performance were differentiated. However, the impact of openness on predictions was limited when compared to other predictors (measured through permutation importance in Study I and Study II) and openness showed no relation to learning in some contexts (i.e., predicting grades in Study I and performance on PCs in Study II).

In additional analyses for the third study of this dissertation a significant link between affective SR processes and neuroticism was found (see Appendix B). Specifically, the group of participants that was characterized by primarily negative emotions, also showed significantly higher values of neuroticism than other participants but no differences in other personality facets. Together with the result that this group of participants with high levels of negative emotions learned significantly less than other learners (see Study III) this is in line with previous findings of potential detrimental effects of neuroticism on academic performance (Chamorro-Premuzic & Furnham, 2003). However, no direct link between neuroticism and learning outcomes was identified in this dissertation. Instead, the findings indicate that the negative emotional processes during might have mediated the effect of neuroticism on learning, which would support the general assumption that SR processes mediate the effect of personal dispositions on learning outcomes (Schunk & Greene, 2018a).

Overall, relations between personal dispositions and learning outcomes were found in the present dissertation, but their impact was less pronounced than for constructs from other areas of the proposed framework (i.e., learning activities, driving forces, and limited resources). A potential explanation of these findings lies in differences between state and trait levels of regulation. Specifically, research suggest that individual with higher levels of trait self-control engage in less in-situ regulation of behaviors (Hill et al., 2014; Hofmann, Baumeister, Förster, & Vohs, 2012; Inzlicht et al., 2021). Instead of engaging in SR in the moment these individuals find alternative ways to regulate (e.g., by avoiding specific situations that require high levels of SR). However, such strategies are not applicable in laboratory learning tasks which is a potential explanation for the low impact of trait measures of SR in the studies of this dissertation. Further does the proposed framework posit that the effect of personal dispositions on learning outcomes and achievement is mediated over the interaction of SR processes that unfold in the learning session. Indications for this assumption were only tested in additional analyses of Study III. Other methodological approaches might therefore be necessary to reveal more pronounced effects of personal dispositions.

All in all, personal dispositions in the present study showed less pronounced effects on learning outcomes across contexts than other research traditions on SR in education. However, the limited predictive value in the present dissertation is potentially related to the methodological approach and low association between trait measures of SR and SR processes in specific situations (Inzlicht et al., 2021). Nonetheless, personal dispositions showed value for an integrative framework on SR in education in specific areas of each study of this dissertation.

5.1.2.4 Limited resources

The regulatory processes comprised in SR require different kinds of cognitive processing resources to be initiated, maintained, and adjusted. These processes most prominently include EF and working memory. The studies of the present dissertation provided empirical evidence supporting their importance for learning outcomes in an integrative framework of SR. Working memory (i.e., d' scores for a two-back and three-back task) showed high predictive value across different outcomes in Study I. Particularly the d' score of a three back task was consistently included in predictions of laboratory learning task performance and among the most frequently selected predictors for grades. This measure further provided the largest predictive value in laboratory learning tasks and positive contributions in predictions of grades. The equivalent measures for a two-back task showed a similar pattern and therefore underlined the importance of working memory in both contexts (i.e., school and laboratory learning tasks). These findings directly correspond to empirical findings on the role working memory and EF in education. As many studies have shown, working memory represents a cognitive resource that has shown strong associations with learning and academic achievement across many educational contexts (Alloway, 2006; Alloway & Alloway, 2010; Friso-van den Bos, van der Ven, Kroesbergen, & van Luit, 2013). Available working memory resources seem to be task dependent (Turner & Engle, 1989), especially when working memories effect on learning is investigated via the imposed cognitive load in a given learning situation (Castro-Alonso et al., 2021; Paas & Sweller, 2012; Sweller, 1994). Yet, the findings of Study I suggest that working memory has a strong, context-independent relation to learning. In other words, working memory seems to be an underlying cognitive resource that is beneficial regardless of the learning situation. Even though, the experienced cognitive load significantly depends on characteristics of the learning task and environment (e.g., the design of the learning material, Castro-Alonso et al., 2021) individuals with higher levels working memory resources will likely still have more cognitive resources at their disposal in high load-inducing situations than individuals with lower initial levels of working memory resources. These resources are required to carry out self-regulatory processes. This assumption is further supported by research on working memory interventions, which have shown great promise in fostering academic achievement across academic domains as well as for low and high performing students (e.g., Titz & Karbach, 2014).

Results of the second study, on the other hand, found no relations between working memory and learning outcomes. Instead switching has demonstrated great predictive value for learning when using tablets. In contrast to findings of Study I, this demonstrates that the most important cognitive resource is task dependent. The art learning environment used in Study II consisted of two main components (i.e., the learning and action panel). To use this environment to its full potential, participants were required to integrate information between these panels. Specifically, they were required to shift their attention between exploring the art works displayed in the action panel and the guiding questions provided in the learning panel. The ability to switch attention between tasks is the core concept of task switching (Miyake et al., 2000). From an instructional point of view this finding is in line with previous studies in similar learning environments, which have shown that specific cognitive resources are often required for learners to benefit from complex learning environments (Kornmann, Kammerer, Anjewierden, et al., 2016; Kornmann, Kammerer, Zettler, et al., 2016). This underlines the notion of the proposed framework that the interplay of cognitive resources and other self-regulatory processes unfolds within a given learning situation.

A noteworthy finding across Study I and Study II is that the importance of cognitive resources to predict learning increased in more extreme comparison. This indicates that both working memory and switching do not exclusively represent minimum requirements to effectively engage in learning task. Instead, very high levels of working memory and EFs enabled learners to achieve superior performance in the learning tasks. This assumption was further supported, by differential pattern of the two working memory measures in Study I. In this study the ‘easier’ working memory task (i.e., two-back task) was important for prediction at less extreme cutoffs whereas the ‘harder’ working memory task (i.e., three-back task) was particularly influential in more extreme cutoffs. This indicates that one task measured the minimum working memory requirements necessary to effectively work with the materials, while the other represented a potential boosting factor that can enable the highest levels of achievement. Similar patterns can be derived from working memory intervention studies, that are more efficient for learners with deficits but can also benefit learners with high abilities (Titz & Karbach, 2014).

Taken together, the studies of the present dissertation provided strong empirical support for the essential role of cognitive resources in SR processes. The importance of cognitive resources was shown both in a context-independent fashion for working memory and learning task specific manner for task switching. This underlines that an integrative approach to SR in education needs to consider cognitive resources as a key factor, but that the relative importance of certain resources (e.g., task switching) depends on characteristics of the learning task and environment. To further put the findings outlined in the previous sections into perspective, in the following paragraphs strength and limitations of the studies in the present dissertation will be discussed.

5.2 Strength and limitations

The studies of the present dissertation provide initial empirical evidence for an integrative framework of SR in education. Consolidating research on learning activities, driving forces, personal dispositions, and limited resources into this framework required a comprehensive set of measures collected across all studies and state of the art methodological approaches. These elements represent strength of the present dissertation, but also encompass challenges and limitations and will be discussed in the next sections.

5.2.1 Measures

The vastly differing approaches related to SR in education are also reflected in many different ways to measures self-regulatory constructs. A major strength of the present dissertation is that all three studies incorporated measures of self-regulated learning at different levels of granularity. Study I and Study II used a comprehensive data set that included key measurement instruments from all areas of SR research in education investigated in this dissertation. These encompassed self-reports of strategy use (e.g., rehearsal, critical evaluation), questionnaires regarding affect and motivation and their regulation (e.g., expectancy value, emotion regulation strategies), personal dispositions (e.g., personality and grit), and measures of working memory and executive functions obtained from reaction times and error rates in cognitive task (e.g., n-back, task shifting paradigm). Further, in the second study task specific measures of the learning process, such as prior knowledge and metacognitive judgments made during learning were included. The final study aimed at an even more fine-granular level, by investigating the temporal development of multiple emotions during a learning process and their relation to habitual emotion regulation and personality. However, while the breath of measures incorporated in the present study offered great insights on multiple aspects of SR, it also comprised limitations with regards to the measurement of SRL process. These include the lack of process measures used to capture learning activities and the overreliance on state-like measures of driving-forces and limited resources. Including such processes measures in analyses was far beyond the scope of this dissertation. Nonetheless, this limitation will be discussed in the following section.

Over the years the focus of SRL measures has increasingly shifted towards online process measures (Boekaerts & Cascallar, 2006). Particularly in technology-based learning environments the use of online trace-data to capture learning activities has shown great advances (e.g., think aloud protocols, log file data from hyper-media learning environments, Azevedo et al., 2013). This development is based in and has further contributed to criticisms of the use of self-report measures to capture SRL (e.g., Winne & Perry, 2000). The core issue in these critiques entails that self-reports measure dispositional representations of SRL rather than SRL as processes, that is unfolding during a learning episode. The measurements of learning activities in the first two studies of this dissertation fall subject to this criticism as exclusively self-report measures were used to capture SRL strategies. Therefore, the scope of learning activities in these studies represents what students typically do during their studies and in learning situations, but may not reflect the strategies they actually deployed in the laboratory learning tasks or in school. Further, these questionnaires were not domain- or task-specific. Students preferred learning strategies are likely to differ depending on the domain and task at hand, as the effectiveness of learning strategies is dependent on the content domain (Alexander et al., 2011; Greene et al., 2015). Nevertheless, the learning activities captured in this dissertation provided predictive value in both school (Study I) and laboratory task settings (Study I + Study II) indicating that these surveys assess aspects of SRL that are relevant for learning in different contexts. Further, in an additional study this shortcoming was directly addressed (Freed et al., under review). Specifically, within the art-learning task (see Study II) we tested if navigational learning behaviors obtained from logfiles moderated the effects of cognitive resources (i.e., visuospatial ability) and personal dispositions (i.e., conscientiousness, grit). Results showed that sequences of learning behaviors indicative of more thorough processing of the materials were linked to learning and partially moderated the impact of cognitive resources on learning outcomes. The included personal dispositions showed no relation to leaning or learning activities in this study. Overall, the predictive value of process measures in this study was smaller than the direct effects of prior knowledge and visuospatial ability. This indicated while the inclusion of process measures of learning activities can provide additional insights to the studies of the present dissertation, their impact on the main findings of this work might be limited and further research is required to understand their interplay with the variables researched in this dissertation.

Similar to the issues regarding process measures of learning activities, the way driving forces and particularly limited resources were assessed in this dissertation pose limitations. Limited resources represent underlying cognitive processes required to carry out self-regulatory processes. These processes are highly context dependent and can be affected by other SR processes. For instance, affective processes can bind working memory resources (Kensinger & Corkin, 2003) and the design of learning environments and tasks can alter the cognitive load during learning (Castro-Alonso et al., 2021; Van Merrienboer & Sweller, 2005). Accordingly, to assess the dynamic interplay of limited resources and other areas of SR in a given learning situation, fine granular measurement of available cognitive resources is required. Like in research on online measures of learning activities, the development assessments of limited resources in the field (e.g., Bugl, Schmid, & Gawrilow, 2015) and non-intrusive measures has gained significant attention (e.g., Appel et al., 2019; Sevcenko, Ninaus, Wortha, Moeller, & Gerjets, 2021). Despite these advances and the potential of in-situ measures of cognitive load, the use of such measures was beyond the scope of the present dissertation. Due to the varying nature of learning task considered in the studies of this dissertation, ranging from reading tasks on biological topics (Study I) to the use of complex hypermedia-learning environments (Study II and Study III), obtaining comparable measures of limited resources this dissertation would require further ground laying work that extends the scope of this work. Therefore, measures for limited resources were obtained from specific laboratory tasks (e.g., n-back tasks) at a set time-point during the experiments. While the implied stability of cognitive resources that comes with such measurements does not depict their dynamic nature (Eisenberg et al., 2019), it still provided a sufficiently reliable measures of cognitive resources that showed promising predictive value across learning tasks and contexts (Study I and Study II). Lastly a similar critique can be applied to the measurement of driving forces. In the first two studies of this dissertation these were captured through self-reports of emotion regulation and domain-specific motivation, which does not account for the dynamic nature of such processes (D'Mello & Graesser, 2012). However, in the case of emotions this issue was directly addressed in the third study of this dissertation where emotions as processes were assessed during a learning activity and their relation to learning and personal dispositions was tested.

In sum, the extensive set of measures used is a major strength of the present dissertation. To attain the goal of empirically integrating lines of research on SR in education, comprehensive datasets with measures from multiple research traditions at varying levels of granularity were analyzed. Still the present dissertation did not capture the full breadth of potential variables of interest, specifically with regards to online process measures. Discussing all potential measures that could complement the self-regulatory variables of this dissertation is beyond the scope of this discussion. Prominent self-regulatory constructs that were not included in the studies of this dissertation include achievement goals (Elliot & McGregor, 2001; Elliot & Thrash, 2001), amount of invested mental effort (Salomon, 1984; Schwab, Hennighausen, Adler, & Carolus, 2018), cognitive engagement (Corno & Mandinach, 1983), growth mindsets (Yeager et al., 2019), measures of cognitive load (Paas & Sweller, 2012; Sweller, 1994), measures of learning strategy use (Pintrich, Smith, Garcia, & McKeachie, 1993), metacognitive strategy knowledge (Pintrich, 2002), online measures of learning activities (Azevedo et al., 2013; Greene et al., 2015), resource notions of self-control (Baumeister, Vohs, & Tice, 2007), self-determination (Deci & Ryan, 2012), visual working memory (Luck & Vogel, 1997), and volitional control (McCann & Turner, 2004).

Further strength and limitations are related to methodological approach of the present dissertation, which I will further discuss in the next section.

5.2.2 Methodological approach

The studies in this dissertation were all conducted in controlled laboratory-based settings. Through high degrees of standardization this ensured that results were comparable across studies. Further, to ensure the validity of the overall finding of this dissertation multiple learning outcome measures were investigated. The first study, included self-reported grades from five academic domains as well as the average performance across five laboratory learning tasks. The laboratory learning performance covered a variety of learning measures typically included in laboratory learning settings, such as knowledge questions, inference questions, or items regarding the detection of contradictions in learning materials. Study II and Study III measured learning through domain specific knowledge and inference questions designed by domain experts. Through the variety of these measures the present dissertation can be directly linked to typical outcomes in the respective literatures. Studies on personal dispositions, for instance, typically use academic achievement as an outcome (e.g., GPA or performance in university courses) to establish the importance trait SR constructs in education (e.g., Duckworth et al., 2010; Poropat, 2009). Learning activities in the form of SRL processes, on the other hand, are often researched in laboratory settings where task specific post-test measures are used to measure learning (Azevedo et al., 2013; Zheng, 2016). Yet, integrations of both outcomes in a single sample remain scarce. This issue was directly addressed in the first study of this dissertation were measures of academic achievement and performance in laboratory learning tasks were investigated in one sample. However, the measures of academic achievement was retrospective and self-reported, which requires cautious interpretation of the results (Kuncel, Credé, & Thomas, 2005).

While the studies in this dissertation investigated the importance of SR constructs for learning outcomes in controlled laboratory settings, it is important to note that all studies were cross-sectional and correlational. This design approach is in line with the objective of this study to extend the scope of research on SRL in educational settings but it is not suitable to reveal causal relationships between the constructs investigated in the three studies of this dissertation. Particularly the results regarding the prediction of grades in the first study have to be cautiously interpreted in terms of causality. Since grades were collected retrospectively the possibility that they had a causal impact on other constructs remains. Such effects of grades on self-efficacy, goals, and motivation have been previously established (e.g., Shim & Ryan, 2005). Nevertheless, given that the present dissertation aimed to provide first empirical evidence for an integrative framework, the correlational study design was necessary and suitable. Detailed causal understanding of the relations between specific learning activities, driving forces, personal dispositions, and limited resources and their development require extensive longitudinal research programs that are far beyond the scope of this dissertation (see chapter 5.3.3).

The research question addressed in this study by design required a large number of potential independent variables and multiple potential outcomes (e.g., grades and laboratory task performance, see Study I) to be jointly analyzed. To circumvent potential issues of classical statistical analyses in this context (e.g., false positive results, overfitted models; Dwyer et al., 2018; Eklund et al., 2016; Whelan & Garavan, 2014) robust and explainable modelling procedures were used in all studies of this dissertation. In the first two studies, ensemble machine learning predictions were conducted to extract the most robust predictors of learning outcomes in different context. Furthermore, parsimonious explainable models were selected over more complex, opaque models to ensure interpretable results. However, these models are limited in their capability to reveal interactions and complex, non-linear trends. The lack of analyses on the interplay between predictors is a limitation of this dissertation. Due to the large number of variables included in analyses, exploring possible interactions was not feasible. For instance, accounting for all possible two-way interactions of the 39 predictors in the first study would have resulted in 1482 features for analyses. Therefore, potential interactions should be investigated in targeted analyses in future studies. As outlined in the framework, interactions such as potential effects of driving forces on cognitive resources and vice versa (e.g., Kensinger & Corkin, 2003), are an essential part of self-regulatory processes in learning situations. The final study of this dissertation used robust statistical modelling approaches to depict the relation between multiple emotions and learning. Moreover, it provided indirect support for the assumption of the proposed framework that the effect of personal dispositions on learning is mediated through learning activities. Specifically, additional analyses revealed that the group of students with the most negative emotional experience was characterized significantly higher levels of neuroticism the neutral and positive groups of students (see Appendix B). Given that the personality measures in this study were collected on the day before the learning activity this indicates that the trait neuroticism was a causing factor or catalyst for the negative emotions these students experienced, which ultimately led to lower learning gains. However, direct tests of interactions between parts of the proposed model of this dissertation and mediational analyses are required in future research to unravel potential compensatory effects and causal relations.

All in all, the present study used state of the art methodology to find initial support for the proposed framework. Still further research using different designs and methodological approaches is needed to provide empirical support for the mechanisms of this integrative model of SRL. The next chapter will briefly outline potential avenues for future research to achieve this goal, before practical implications of this work are discussed.

5.3 Future directions and implications

Through the integration of research on learning activities, driving forces, personal dispositions, and limited resources into a framework of SR in education this dissertation aimed to provide a common ground that different research traditions related to SR in education can build upon. This broader perspective on SR was built around learning activities, which were derived from a synthesis of models on SRL (Panadero, 2017; Puustinen & Pulkkinen, 2001). Subsequently, this core concept of SR was enhanced by adjacent fields of research that showed great conceptual relations to learning activities but demonstrated varying levels of empirical integration with SRL. By demonstrating how constructs from the four areas of research on SR jointly predicted learning outcomes at different levels of granularity (i.e., learning in school and laboratory learning tasks) this dissertation empirically started to bridge the gap between different approaches to investigate SR in education. Outlining the potential interrelations between learning activities, driving forces, personal dispositions, and cognitive resources is a major contribution of this work to the field of research on SR. The implications of this integrative approach to SR in education and potential future directions will be discussed below. First, key implications and resulting directions for future research will be discussed. Second, implications for practitioners will be outlined.

5.3.1 Implications for research

Three central implications and corresponding future directions will be discussed in the following sections. First, the value of incorporating cognitive resources into research on SRL will be outlined. Second, the importance of potential interactions and compensatory effects between the core areas of the proposed framework are detailed. Lastly, further steps towards consolidating approaches on SR in education will be deliberated.

5.3.1.1 Limited resources in SRL research

One of the most striking features that arose while interconnections between adjacent fields of research on SR were identified was the lack of integration of cognitive resources in research on SRL. Except for one theory of SRL (i.e., Borkowski et al., 2000) EF and working memory are sparsely incorporated in theories and models of SRL. This is in direct contrast with the cognitively demanding nature of SRL (e.g., Schunk & Greene, 2018b) and the ample amount of empirical support that links cognitive resources to learning and academic achievement (Alloway & Alloway, 2010; Titz & Karbach, 2014). While initial attempts to bridge the gap between research on learning activities and limited resources begin to arise (de Bruin & van Merriënboer, 2017; Effeney et al., 2013), a systematic integration of these two research traditions is still far off. The results of the studies in the present dissertation further emphasize this issue. Cognitive resources were key predictors of performance across all contexts investigated in the present dissertation. The patterns of results indicated that limited resources can be a minimum requirement to effectively engage in certain learning activities or boosting factors that enable high levels performance. The importance of sufficient cognitive resources to make use of complex learning environments (Kornmann, Kammerer, Anjewierden, et al., 2016) can directly inform research on SRL in computer-based learning environments. In this line of research ways to overcome student's overreliance on ineffective, surface level strategies are developed and investigated (e.g., Azevedo, 2005; Narciss et al., 2007). The frequent use of simple and ineffective learning strategies itself can be caused by a (temporal) lack of cognitive resources. Studies have shown that complex hypermedia-learning environments can foster learning but also imposes additional self-regulatory demands on learners (Opfermann, Scheiter, Gerjets, & Schmeck, 2013). Further, relations between cognitive load and the use of effective learning strategies have been shown (Scheiter, Gerjets, Vollmann, & Catrambone, 2009). This indicates that the deployment of learning activities is dependent on underlying cognitive resources. The core processes of SRL – metacognitive monitoring and control – have also shown to be affected by cognitive resources (Follmer & Sperling, 2016). Future research on SRL should therefore incorporate cognitive resources and identify how they affect learning activities.

Limited resources also offer a potential explanation to one of the major findings of Study III – the predominant negative effect of negative emotions and the lack of beneficial effects of positive emotions. While negative emotions can tax working memory capacity (Curci, Lanciano, Soleti, & Rimé, 2013) and therefore reduce available cognitive resources necessary for SR. Positive emotions on the other hand cannot enhance working memory beyond its regular capabilities. Based on the third study of this dissertation future studies on affective SR processes during learning should test if and how negative affective experiences show detrimental effects of cognitive resources during learning.

All in all, the examples outlined above demonstrate the great potential that cognitive resources have to inform research on SRL. Future studies should build upon this notion and investigate differences in SR in relation to underlying cognitive resources.

5.3.1.2 Interactions and compensatory effects

A key implication of this dissertation for future research is the extension of the scope of research on SR in education. The studies in this dissertation have shown how conceptualizing SR as the interaction of learning activities, driving forces, personal dispositions, and cognitive resources paints a more complete picture of self-regulatory processes in educational settings. Future research should build upon these initial findings by directly investigating interactions between learning activities, driving forces, personal dispositions, and limited resources. Potential interaction and compensation effects between these areas of research are manifold. For instance, previous studies have shown that interest and conscientiousness show a distinct interaction pattern (Trautwein et al., 2015). Specifically, conscientiousness is only crucial in determining academic effort when students are lacking interest in a subject. Related interaction effects are implied by models of SRL, where SRL skills only lead to desirable learning outcomes if students are motivated to engage in the task (Efklides, 2011; Zimmerman, 2000). In the previous section I have outlined that similar effects can occur with regards to underlying cognitive resources. For example, students' knowledge of adequate learning activities and motivation to deploy them will only show positive impact on learning if the sufficient cognitive resources are available. In addition to interaction effects, compensatory effects are also likely to occur during SR processes. For instance, when no strict time constrains are imposed, a lack in the use of appropriate learning activities might be compensated through personality and driving forces (e.g., repeatedly going over the material). An example for both a potential interaction or compensatory effect that is commonly found in SRL research, but not extensively discussed is the role of prior knowledge. Prior knowledge is a crucial predictor of learning outcomes and further can directly impact the SRL processes in multiple ways. The studies in this dissertation that included prior knowledge (i.e., Study II and Study III) showed that prior knowledge is a strong predictor of learning outcomes in different tasks. Research suggests that the impact of prior knowledge on SRL goes even further. Specifically, knowledge can predict if, how, and which cognitive and metacognitive strategies are engaged in during the learning process (Taub & Azevedo, 2019; Taub, Azevedo, Bouchet, & Khosravifar, 2014). On the other hand, sufficient prior knowledge can serve as 'short cut' in the SR processes. Borkowski et al. (2000) proposed that given sufficient domain-specific knowledge to solve the current task or problem, learners can skip most of the regulatory activities and directly solve the issue. In this case little strategic planning and monitoring are required and the focus shifts mainly on the application of knowledge akin to schemas that are executed (Norman & Shallice, 1986; Sweller, 1994). Yet, even in such cases certain levels of regulation are required. As pointed out for strategy use, without the motivation to apply learning activities, prior knowledge will not suffice to solve a task (Zimmerman, 2000).

In sum, previous research and the proposed framework highlight that learning activities, driving forces, personal dispositions, and limited resources are interrelated parts of SR. Further research should focus on revealing and confirming interaction and compensatory patterns.

5.3.1.3 Steps towards an integrative framework of SRL

As outlined in the methodological strength and limitations, the studies of the present dissertation by the nature of their questions and design do not warrant definitive causal claims. However, in order to achieve a comprehensive understanding of SR in education an in-depth understanding of the causal relations between learning activities, driving forces, personal dispositions and limited resources is imperative. A promising approach to pursue this goal could be based on research on cognitive resources. In this line of research, the development of key EF and working memory is investigated in early developmental stages (e.g., Cowan & Alloway, 2009). Studies in this area of research have shown that cognitive resources are crucial for the individual development, learning, and later academic achievement (Alloway, 2006; Cowan, 2014; Nayfeld, Fuccillo, & Greenfield, 2013). Cognitive resources are further closely related to (everyday) SR (Hofmann, Friese, Schmeichel, & Baddeley, 2011; Hofmann, Schmeichel, et al., 2012; Ilkowska & Engle, 2010). A key step for research on SR education lies in the combination of these approaches with other facets of SR in education (e.g., learning activities, personal dispositions, and driving forces). Basing models on the acquisition of SRL strategies (Borkowski et al., 2000) in developmental patterns of cognitive resources could be a fruitful step to understand how (meta-)cognitive strategy use develops and through which learning activities individual differences in the development affect academic outcomes. In this context, early deficits or advantages in cognitive resources could explain if and how students internalize the use of specific learning strategies, which builds the basis for the later development of SR skills (Borkowski et al., 2000). For example, student's repertoire of learning strategies might differ because more elaborative strategies might not be commonly used if a lack of cognitive resources prevents that these learning activities can be effectively used. These connections between cognitive resources and learning activities can further be related to the development of driving forces. For instance, the lack of success during the acquisition of learning activities due to insufficient cognitive resources can lead to negative affective experiences (e.g., through lower levels of progress than expected, Carver & Scheier, 1998). Such negative driving forces in turn can lead to negative affective and motivational associations with the specific learning strategy, the task, and content domain, which can impede future performance in similar tasks and situations. Similar to the examples related to cognitive resources outlined above, investigations on causal relation between learning activities, driving forces, personal dispositions, and limited resources can further be grounded in developmental accounts of personality (Caspi, Roberts, & Shiner, 2004), self-control (Moffitt et al., 2011) or motivation (Wigfield et al., 2015).

Taken together, the present dissertation provided initial evidence for an integrative framework of SR in education. Future research should build upon these findings by investigating causal relations and developmental trajectories of the key areas of the proposed framework.

5.3.2 Practical implications

Beyond its contribution to the scientific debate on SR in education this dissertation further has key implications for practitioners. The central role of SR to face key challenges in concurrent and future education is well established in both scientific and practical contexts (OECD, 2013, 2018; The World Bank Group, 2011). The proposed framework provides practitioners with an overview of the core process of

SRL. Specifically, based on previous synthesis of theories and models of SRL (Panadero, 2017; Puustinen & Pulkkinen, 2001) a concise core cycle of SR has been outlined. This process depicts the essence of theories and models of SRL and includes a preparation, a performance, and an appraisal phase. Practitioners aiming to foster SR in students should focus on this core process when incorporating SR in their teaching. As outlined by models that focus on the development of SRL (Borkowski et al., 2000) strategies need to acquired, applied, and internalized before students can gain the ability to adaptively regulate their learning through strategy selection and adaption. The core process of SRL outlined in this dissertation provides a simple and effective structure that can be incorporated in learning situations to foster the acquisition of SR skills. However, the main contribution of the present dissertation lies in the extension of these learning activities by driving forces, personal dispositions, and limited resources. This functional differentiation of constructs related to SR with different underlying mechanism can serve practitioners as a landscape to situate research in and plan interventions targeting different areas of SR in education. The complex, interconnected nature of the core areas of the proposed framework, together with empirical findings provided through the studies of this dissertation can be a boon or bane for practitioners.

Key points for practitioners in this context are the potential interactions and compensatory effects that are encapsulated in SR. Compensatory mechanisms highlight that disadvantageous personal dispositions (e.g., low grit or conscientiousness) can be compensated in multiple ways, for instance, through high levels of driving forces (e.g., interest Trautwein et al., 2015). On the other hand, interaction effects can also be troublesome for practitioners, when newly acquired SR skills may not translate to better learning outcomes because students lack motivation (Zimmerman, 2000) or the cognitive resources to effectively use learning materials (e.g., Kornmann, Kammerer, Anjewierden, et al., 2016).

Further the present dissertation has shown that high learning outcomes across contexts are based on the same structure of self-regulatory processes (i.e., learning activities, driving forces, personal dispositions, and limited resources). This highlights that practitioners should focus on fostering multiple aspects of the SR process to enable maximum achievement (e.g., Zheng, 2016). However, the present dissertation has also shown that the specific most important determinants of learning are highly dependent on the learning task, context, and environment. For instance, Study II showed that using tablets instead of PCs can be associated with significantly higher SR requirements. This indicates that all properties of the learning task need considered to identify how learners can be successful in a given learning situation.

All in all, the present dissertation provides value for practical applications by providing a comprehensive landscape of research on SR in education. It has further shown that the underlying structure of SR is applicable across different domains and levels of granularity. However, the studies of this dissertation further highlighted that the most effective predictors of learning are largely context specific, which indicates that 'one size fits all' solutions with regards to SR in education are unlikely to succeed. In other words, while SR consists of processes that ideally enable learners to succeed in any learning situation, contextualized and potentially individualized approaches to foster learning are likely required to ensure ideal learning outcomes for all students.

5.4 Conclusion

Through technological changes our educational systems have become more challenging and complex. Recent events, particularly the ongoing COVID-19 pandemic, have further increased the need for learners to engage in SRL. In three studies, this dissertation extended the scope of SRL researcher and integrated several research traditions related to SR in educational settings. Specifically, the importance of learning activities, driving forces, personal disposition, and limited resources for different learning outcomes has been investigated. In these studies, robust machine learning and statistical approaches novel to research in SRL have been carried out to obtain reliable empirical evidence for an integrative framework of SR in education. Specifically, drawing from a unique data set the first two studies have investigated the joint predictive value of key variables from four research traditions on SR in education. Study I, showed that predictors from all four main components of the proposed framework substantially predicted school grades and laboratory task performance. The specific predictors partially varied with measures of motivation and working memory capacity being important for both outcomes, while the predictive value for other measures such as the use of rehearsal strategies or effort was specific to one of the outcomes. The second study found that the pattern of predictors revealed in the first study can also be found in predictions of a specific learning task. However, the specific predictors varied. More importantly, this study revealed that self-regulatory requirements are higher when tablets are used for a complex, exploratory art-learning task. The final study showed how multiple emotions unfold throughout a learning session and how learners complex, emotional experiences are related to learning. It showed the unfolding of multiple negative emotions, which is determined by individuals’ affective dispositions, is detrimental to learning. Together, the empirical investigations of this dissertation have shown that integrating SRL into a larger framework of SR in education is essential to obtain a comprehensive understating of the regulation of learning process. Through these results this dissertation made first promising steps towards an integration of SR in education. To achieve this goal extensive research programs on SR education are required.

References

Abrams, R. A., Weidler, B. J., & Suh, J. (2015). Embodied seeing: The space near the hands. In B. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 63, pp. 141-172). Amsterdam, The Netherlands: Elsevier. doi: https://doi.org/10.1016/bs.plm.2015.03.005

Alexander, P. A., Dinsmore, D. L., Parkinson, M. M., & Winters, F. I. (2011). Self- regulated learning in academic domains. In B. Zimmerman & D. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 393–407). New York: Routledge.

Alloway, T. P. (2006). How does working memory work in the classroom? *Educational Research and Reviews*, *1*(4), 134–139.

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*(1), 20–29.

Appel, T., Sevcenko, N., Wortha, F., Tsarava, K., Moeller, K., Ninaus, M., … Gerjets, P. (2019). Predicting cognitive load in an emergency simulation based on behavioral and physiological measures. In W. Gao, H. M. Ling Men. & M. Turk (Eds.), *2019 International Conference on Multimodal Interaction* (pp. 154–163). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3340555.3353735

Ardila, A. (2008). On the evolutionary origins of executive functions. *Brain and Cognition*, *68*(1), 92–99.

Azevedo, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? The role of self-regulated learning. *Educational Psychologist*, *40*(4), 199–209.

Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology*, *96*(3), 523.

Azevedo, R., Harley, J., Trevors, G., Duffy, M., Feyzi-Behnagh, R., Bouchet, F., & Landis, R. (2013). Using trace data to examine the complex roles of cognitive, metacognitive, and emotional self-regulatory processes during learning with multi-agent systems. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 427–449). Amsterdam, The Netherlands: Springer.

Azevedo, R., Johnson, A., Chauncey, A., & Burkett, C. (2010). Self-regulated learning with MetaTutor:Advancing the science of learning with metacognitive tools. In M. S. Khine & I. M. Saleh (Eds.),NewScience of Learning(pp. 225–247). New York: Springer.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*(1), 1–26. https://doi.org/10.1146/annurev.psych.52.1.1

Bauer, I. M., & Baumeister, R. F. (2011). Self-regulatory strength. In *Handbook of self-regulation: Research, theory, and applications, 2nd ed.* (pp. 64–82). New York, NY, US: Guilford Press.

Baumeister, R. F., Alquist, J. L., & Vohs, K. D. (2015). Illusions of learning: Irrelevant emotions inflate judgments of learning. *Journal of Behavioral Decision Making*, *28*(2), 149–158.

Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science*, *16*(6), 351–355.

Bidjerano, T., & Dai, D. Y. (2007). The relationship between the big-five model of personality and self-regulated learning strategies. *Learning and Individual Differences*, *17*(1), 69–81.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*, 417–444.

Boekaerts, M. (1996). Self-regulated learning at the junction of cognition and motivation. *European Psychologist*, *1*(2), 100–112. https://doi.org/10.1027/1016-9040.1.2.100

Boekaerts, M., & Cascallar, E. (2006). How far have we moved toward the integration of theory and practice in self-regulation? *Educational Psychology Review*, *18*(3), 199–210.

Boekaerts, M., & Niemivirta, M. (2000). Self-regulated learning: Finding a balance between learning goalsand ego-protective goals. In M.Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 417–450). New York: Academic.

Boekaerts, M., & Pekrun, P. R. (2016). Emotion and emotion regulation in academic settings. In L. Corno & E. M. Anderman (Eds.), *Handbook of educational psychology* (pp.76–90). New York, NY: Routledge.

Boerner, S., Seeber, G., Keller, H., & Beinborn, P. (2005). Lernstrategien und lernerfolg im studium. *Zeitschrift Für Entwicklungspsychologie Und Pädagogische Psychologie*, *37*(1), 17–26.

Borkowski, J. G., Chan, L. K. S., & Muthukrishna, N. (2000). A process-oriented model of metacognition: links between motivation and executive functioning. In G. Schraw & J. Impara (Eds.), *Issues in the Measurement of Metacognition*. Lincoln, NE: Buros Institute of Mental Measurements, University of Nebraska.

Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, *27*, 1–13.

Brydges, R., Manzone, J., Shanks, D., Hatala, R., Hamstra, S. J., Zendejas, B., & Cook, D. A. (2015). Self-regulated learning in simulation-based training: a systematic review and meta-analysis. *Medical Education*, *49*(4), 368–378.

Bugl, P., Schmid, J., & Gawrilow, C. (2015). Ambulantes Assessment in der Schule: Den schulischen Alltag erfahrbar machen. *Lernen Und Lernstörungen*, *4*(4), 261–268. https://doi.org/10.1024/2235-0977/a000115

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, *65*(3), 245–281. https://doi.org/10.2307/1170684

Butterfield, E. C., Albertson, L. R., & Johnston, J. C. (1995). On making cognitive theory more general and developmentally pertinent. *Research on Memory Development*, *68*, 73–99.

Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review*, *97*(1), 19–35. https://doi.org/10.1037/0033-295X.97.1.19

Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior*. New York, NY: Cambridge University Press. https://doi.org/10.1017/CBO9781139174794

Caspi, A., Roberts, B. W., & Shiner, R. L. (2004). Personality development: Stability and change. *Annual Review of Psychology*, *56*(1), 453–484. https://doi.org/10.1146/annurev.psych.55.090902.141913

Castro-Alonso, J. C., de Koning, B. B., Fiorella, L., & Paas, F. (2021). Five strategies for optimizing instructional materials: Instructor- and learner-managed cognitive load. *Educational Psychology Review*, 1–29.

Chamorro-Premuzic, T., & Furnham, A. (2003). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of Research in Personality*, *37*(4), 319–338.

Cirino, P. T., & Willcutt, E. G. (2017). An introduction to the special issue: Contributions of executive function to academic skills. *Journal of Learning Disabilities*, *50*(4), 355–358.

Corno, L., & Mandinach, E. B. (1983). The role of cognitive engagement in classroom learning and motivation. *Educational Psychologist*, *18*(2), 88–108.

Costa, P. T., & McCrae, R. R. (1998). Six approaches to the explication of facet-level traits: examples from conscientiousness. *European Journal of Personality*, *12*(2), 117–134.

Costa, P. T., & McCrae, R. R. (2008). The revised NEO personality inventory (NEO-PI-R). In G. J. Boyle, G. Matthews & D. H. Saklofske (Eds), *The SAGE handbook of personality theory and assessment, Vol 2: Personality measurement and testing.* (pp. 179–198). Thousand Oaks, CA, US: Sage Publications, Inc. https://doi.org/10.4135/9781849200479.n9

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, *26*, 197–223. https://doi.org/10.1007/s10648-013-9246-y

Cowan, N., & Alloway, T. (2009). Development of working memory in childhood. In M. L. Courage & N. Cowan (Eds.), *The development of memory in infancy and childhood* (pp. 303–342). Hove, East Sussex, UK: Psychology Press.

Curci, A., Lanciano, T., Soleti, E., & Rimé, B. (2013). Negative emotional experiences arouse rumination and affect working memory capacity. *Emotion*, *13*(5), 867-880. https://doi.org/10.1037/a0032492

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, *22*(2), 145–157.

D'Mello, S. K., & Graesser, A. C. (2013). AutoTutor and affective AutoTutor: learning by talking withcognitively and emotionally intelligent computers that talk back. *ACM Transactions on InteractiveIntelligent Systems*, *2*, 1–39. https://doi.org/10.1145/2395123.2395128

Đambić, G., Krajcar, M., & Bele, D. (2016). Machine learning model for early detection of higher education students that need additional attention in introductory programming courses. *International Journal of Digital Technology & Economy*, *1*(1), 1–11.

Dang, J., King, K. M., & Inzlicht, M. (2020). Why are self-report and behavioral measures weakly correlated? *Trends in Cognitive Sciences*, *24*(4), 267–269.

de Boer, H., Donker-Bergstra, A. S., Kostons, D., Korpershoek, H., & van der Werf, M. P. C. (2012). *Effective Strategies for Self-regulated Learning: A Meta-analysis*. Groningen: GION onderzoek/onderwijs.

de Bruin, A. B. H., & van Merriënboer, J. J. G. (2017). Bridging cognitive load and self-regulated learning research: A complementary approach to contemporary issues in educational research. *Learning and Instruction*, *51*, 1–9.

De Raad, B., & Schouwenburg, H. C. (1996). Personality in learning and education: A review. *European Journal of Personality*, *10*(5), 303–336. https://doi.org/10.1002/(SICI)1099-0984(199612)10:5<303::AID-PER262>3.0.CO;2-2

Deci, E. L., & Ryan, R. M. (2012). Self-determination theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (Vol. 1, pp. 416–437). Thousand Oaks, CA: Sage. https://doi.org/10.4135/9781446249215.n21

Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review*, *28*(3), 425–474. https://doi.org/10.1007/s10648-015-9320-8

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168.

Dignath, C., Buettner, G., & Langfeldt, H.-P. (2008). How can primary school students learn self-regulated learning strategies most effectively?: A meta-analysis on self-regulation training programmes. *Educational Research Review*, *3*(2), 101–129.

Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, *3*(3), 231–264.

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The Mini-IPIP Scales: Tiny-yet-effective measures of the Big Five Factors of Personality. *Psychological Assessment*, *18*, 192–203. https://doi.org/10.1037/1040-3590.18.2.192

Dörrenbächer, L., & Perels, F. (2016). Self-regulated learning profiles in college students: Their relationship to achievement, personality, and the effectiveness of an intervention to foster self-regulated learning. *Learning and Individual Differences*, *51*, 229–241.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, *92*, 1087–1101. https://doi.org/10.1037/0022-3514.92.6.1087

Duckworth, A. L., Tsukayama, E., & May, H. (2010). Establishing causality using longitudinal hierarchical linear modeling: An illustration predicting achievement from self-control. *Social Psychological and Personality Science*, *1*(4), 311–317.

Duffy, M. C., & Azevedo, R. (2015). Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system. *Computers in Human Behavior*, *52*, 338–348.

Dwyer, D. B., Falkai, P., & Koutsouleris, N. (2018). Machine learning approaches for clinical psychology and psychiatry. *Annual Review of Clinical Psychology*, *14*, 91–118.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*(1), 109–132.

Effeney, G., Carroll, A., & Bahr, N. (2013). Self-regulated learning and executive function: exploring the relationships in a sample of adolescent males. *Educational Psychology*, *33*(7), 773–796.

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist*, *46*(1), 6–25. https://doi.org/10.1080/00461520.2011.538645

Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes*, *105*(1), 98–121.

Eilam, B., Zeidner, M., & Aharon, I. (2009). Student conscientiousness, self-regulated learning, and science achievement: An explorative field study. *Psychology in the Schools*, *46*(5), 420–432.

Eisenberg, I. W., Bissett, P. G., Enkavi, A. Z., Li, J., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Uncovering the structure of self-regulation through data-driven ontology discovery. *Nature Communications*, *10*(1), 1–13.

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, *113*(28), 7900–7905.

Ekman, P. (1992). Are there basic emotions? *Psychological Review*, *99*, 550–553. https://doi.org/10.1037/0033-295X.99.3.550

Elliot, A. J., & McGregor, H. A. (2001). A 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology*, *80*, 501–519. https://doi.org/10.1037/0022-3514.80.3.501

Elliot, A. J., & Thrash, T. M. (2001). Achievement goals and the hierarchical model of achievement motivation. *Educational Psychology Review*, *13*(2), 139–156.

Field, J. (2000). *Lifelong learning and the new educational order.* Stoke on Trent:Trentham.

Flake, J. K., Barron, K. E., Hulleman, C., McCoach, B. D., & Welsh, M. E. (2015). Measuring cost: The forgotten component of expectancy-value theory. *Contemporary Educational Psychology*, *41*, 232–244. https://doi.org/https://doi.org/10.1016/j.cedpsych.2015.03.002

Follmer, D. J., & Sperling, R. A. (2016). The mediating role of metacognition in the relationship between executive function and self-regulated learning. *British Journal of Educational Psychology*, *86*(4), 559–575.

Fredrickson, B. L., & Branigan, C. (2005). Positive emotions broaden the scope of attention and thought-action repertoires. *Cognition & Emotion*, *19*(3), 313–332.

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186–204.

Friso-van den Bos, I., van der Ven, S. H. G., Kroesbergen, E. H., & van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review*, *10*, 29–44. https://doi.org/https://doi.org/10.1016/j.edurev.2013.05.003

Gatzka, T. (2021). Aspects of openness as predictors of academic achievement. *Personality and Individual Differences*, *170*, 110422. https://doi.org/https://doi.org/10.1016/j.paid.2020.110422

Gatzka, T., & Hell, B. (2018). Openness and postsecondary academic performance: A meta-analysis of facet-, aspect-, and dimension-level correlations. *Journal of Educational Psychology*, *110*(3), 355–377. https://doi.org/10.1037/edu0000194

Greene, J. A., & Azevedo, R. (2007). A theoretical review of Winne and Hadwin's model of self-regulated learning: New perspectives and directions. *Review of Educational Research*, *77*(3), 334–372.

Greene, J. A., & Azevedo, R. (2009). A macro-level analysis of SRL processes and their relations to the acquisition of a sophisticated mental model of a complex system. *Contemporary Educational Psychology*, *34*(1), 18–29.

Greene, J. A., Bolick, C. M., Jackson, W. P., Caprino, A. M., Oswald, C., & McVea, M. (2015). Domain-specificity of self-regulated learning processing in science and history. *Contemporary Educational Psychology*, *42*, 111–128.

Gross, J. J. (2013). Emotion regulation: Taking stock and moving forward. *Emotion*, *13*, 359–365. https://doi.org/10.1037/a0032135

Hacker, D. J., Dunlosky, J., & Graesser, A. C. (2009). *Handbook of metacognition in education*. New York, NY, US: Routledge/Taylor & Francis Group.

Hadwin, A. F., Järvelä, S., & Miller, M. (2011). Self-regulated, co-regulated, and socially shared regulation oflearning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning andperformance* (pp. 65–86). New York: Routledge.

Hadwin, A. F., Järvelä, S., & Miller, M. (2018). Self-regulation, co-regulation and shared regulation in collab-orative learning environments. In D. Schunk & J. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 83–106). New York, NY: Routledge.

Hammond, M., & Collins, R. (2013). *Self-directed learning: Critical practice*. London; UK: Routledge.

Harley, J. M., Pekrun, R., Taxer, J. L., & Gross, J. J. (2019). Emotion regulation in achievement situations: An integrated model. *Educational Psychologist*, *54*(2), 106–126.

Haßler, B., Major, L., & Hennessy, S. (2016). Tablet use in schools: A critical review of the evidence for learning outcomes. *Journal of Computer Assisted Learning*, *32*(2), 139–156.

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A Meta-Analysis. *Review of Educational Research*, *66*(2), 99–136. https://doi.org/10.3102/00346543066002099

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

Hilbert, S., Coors, S., Kraus, E. B., Bischl, B., Frei, M., Lindl, A., … Stachl, C. (2021). *Machine Learning for the Educational Sciences*. PsyArXiv*. March*, *27*.

Hill, P. L., Nickel, L. B., & Roberts, B. W. (2014). Are you in a healthy relationship? Linking conscientiousness to health via implementing and immunizing behaviors. *Journal of Personality*, *82*(6), 485–492.

Hofmann, W., Baumeister, R. F., Förster, G., & Vohs, K. D. (2012). Everyday temptations: An experience sampling study of desire, conflict, and self-control. *Journal of Personality and Social Psychology, 102*(6), 1318–1335. https://doi.org/10.1037/a0026545

Hofmann, W., Friese, M., Schmeichel, B. J., & Baddeley, A. D. (2011). *Working memory and self-regulation.* In K. D. Vohs & R. F. Baumeister (Eds.), *Handbook of self-regulation: Research, theory, and applications* (p. 204–225). Guilford Press.

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

Hoyle, R. H., & Dent, A. L. (2018). *Developmental trajectories of skills and abilities relevant for self-regulation of learning and performance.* In D. H. Schunk & J. A. Greene (Eds.), *Educational psychology handbook series. Handbook of self-regulation of learning and performance* (p. 49–63). Routledge/Taylor & Francis Group.

Huang, C. (2011). Achievement goals and achievement emotions: A meta-analysis. *Educational Psychology Review*, *23*(3), 359. https://doi.org/10.1007/s10648-011-9155-x

Huang, C. (2016). Achievement goals and self-efficacy: A meta-analysis. *Educational Research Review*, *19*, 119–137.

Hutchins, S. D., Wickens, C. D., Carolan, T. F., & Cumming, J. M. (2013). The influence of cognitive load on transfer with error prevention training methods: A meta-analysis. *Human Factors*, *55*(4), 854–874.

Ilkowska, M., & Engle, R. W. (2010). *Working memory capacity and self-regulation.* In R. H. Hoyle (Ed.), *Handbook of personality and self-regulation* (p. 265–290). Wiley-Blackwell. https://doi.org/10.1002/9781444318111.ch12

Inzlicht, M., Werner, K. M., Briskin, J. L., & Roberts, B. W. (2021). Integrating models of self-regulation. *Annual Review of Psychology*, *72*, 319–345.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), e124.

Jansen, M., Lüdtke, O., & Schroeders, U. (2016). Evidence for a positive relation between interest and achievement: Examining between-person and within-person variation in five domains. *Contemporary Educational Psychology*, *46*, 116–127.

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.

Kensinger, E. A., & Corkin, S. (2003). Effect of Negative Emotional Content on Working Memory and Long-Term Memory. *Emotion, 3*(4), 378–393. https://doi.org/10.1037/1528-3542.3.4.378

Koriat, A. (2012). The relationships between monitoring, regulation and performance. *Learning and Instruction*, *22*(4), 296–298.

Kornmann, J., Kammerer, Y., Anjewierden, A., Zettler, I., Trautwein, U., & Gerjets, P. (2016). How children navigate a multiperspective hypermedia environment: The role of spatial working memory capacity. *Computers in Human Behavior*, *55*, 145–158. https://doi.org/10.1016/j.chb.2015.08.054

Kornmann, J., Kammerer, Y., Zettler, I., Trautwein, U., & Gerjets, P. (2016). Hypermedia exploration stimulates multiperspective reasoning in elementary school children with high working memory capacity: A tablet computer study. *Learning and Individual Differences*, *51*, 273–283. https://doi.org/10.1016/j.lindif.2016.08.041

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.

Kuncel, N. R., Credé, M., & Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research*, *75*(1), 63–82.

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

Levine, L. J., & Burgess, S. L. (1997). Beyond general arousal: Effects of specific emotions on memory. *Social Cognition*, *15*(3), 157–181.

Lockhart, R. S., & Craik, F. I. M. (1990). Levels of processing: A retrospective commentary on a framework for memory research. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *44*(1), 87–112. https://doi.org/10.1037/h0084237

Loderer, K., Pekrun, R., & Lester, J. C. (2020). Beyond cold technology: A systematic review and meta-analysis on emotions in technology-based learning environments. *Learning and Instruction*, *70*, 101162. https://doi.org/https://doi.org/10.1016/j.learninstruc.2018.08.002

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281.

Marsh, H. W. (1990). A multidimensional, hierarchical model of self-concept: Theoretical and empirical justification. *Educational Psychology Review*, *2*(2), 77–172. https://doi.org/10.1007/BF01322177

Martin, J., & McLellan, A.-M. (2008). The educational psychology of self-regulation: A conceptual and critical analysis. *Studies in Philosophy and Education*, *27*(6), 433–448.

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does “failure to replicate” really mean? *American Psychologist*, *70*(6), 487.

McCann, E. J., & Turner, J. E. (2004). Increasing student learning through volitional control. *Teachers College Record*, *106*(9), 1695–1714.

McClelland, M., Geldhof, J., Morrison, F., Gestsdóttir, S., Cameron, C., Bowers, E., . . . Grammer,J. (2018). Self-Regulation. In N. Halfon, C. B. Forrest, R. M. Lerner, & E. M. Faustman (Eds.), *Handbook of life course health development* (pp. 275–298). Cham: Springer InternationalPublishing. doi:10.1007/978-3-319-47143-3_12

Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science*, *18*(3), 159–163. https://doi.org/10.1111/j.1467-8721.2009.01628.x

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*(1), 167–202.

Mischel, W., Ayduk, O. N., Berman, M., Casey, B. J., Jonides, J., Kross, E., … Shoda, Y. (2011). “Willpower” over the life span: Decomposing impulse control. *Social Cognitive Affective Neuroscience*, *6*, 252–256.

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., … Ross, S. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, *108*(7), 2693–2698.

Muenks, K., Wigfield, A., Yang, J. S., & O’Neal, C. R. (2017). How true is grit? Assessing its relations to high school and college students’ personality characteristics, self-regulation, engagement, and achievement. *Journal of Educational Psychology*, *109*(5), 599–620. https://doi.org/10.1037/edu0000153

Muis, K. R., & Edwards, O. (2009). Examining the stability of achievement goal orientation. *Contemporary Educational Psychology*, *34*(4), 265–277. https://doi.org/https://doi.org/10.1016/j.cedpsych.2009.06.003

Mulet, J., Van De Leemput, C., & Amadieu, F. (2019). A critical literature review of perceptions of tablets for learning in primary and secondary schools. *Educational Psychology Review*, *31*(3), 631–662.

Nagengast, B., Marsh, H. W., Scalas, L. F., Xu, M. K., Hau, K.-T., & Trautwein, U. (2011). Who Took the “×” out of Expectancy-Value Theory?: A Psychological Mystery, a Substantive-Methodological Synergy, and a Cross-National Generalization. *Psychological Science*, *22*(8), 1058–1066. https://doi.org/10.1177/0956797611415540

Namkung, J. M., Peng, P., & Lin, X. (2019). The relation between mathematics anxiety and mathematics performance among school-aged students: a meta-analysis. *Review of Educational Research*, *89*(3), 459–496.

Narciss, S., Proske, A., & Koerndle, H. (2007). Promoting self-regulated learning in web-based learning environments. *Computers in Human Behavior*, *23*(3), 1126–1144. https://doi.org/https://doi.org/10.1016/j.chb.2006.10.006

National Research Council. (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. Washington, DC: The National Academies Press. https://doi.org/10.17226/13398

Nayfeld, I., Fuccillo, J., & Greenfield, D. B. (2013). Executive functions in early learning: Extending the relationship between executive functions and school readiness to science. *Learning and Individual Differences*, *26*, 81–88. https://doi.org/https://doi.org/10.1016/j.lindif.2013.04.011

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In *Metacognition: Knowing about knowing.* (pp. 1–25). Cambridge, MA, US: The MIT Press.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control ofbehavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.),Consciousness andself-regulation: Advances in research and theory(Vol. 4, pp. 1–18). New York: Plenum.

OECD. (2013). *Trends Shaping Education 2013*. Paris. France: OECD. https://doi.org/10.1787/trends_edu-2013-en

OECD. (2018). *The future of education and skills: Education 2030*. Paris, France: OECD.

Opfermann, M., Scheiter, K., Gerjets, P., & Schmeck, A. (2013). Hypermedia and self- regulation: An interplay in both directions. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 129-141). Amsterdam, The Netherlands: Springer. https://doi.org/10.1007/978-1-4419-5546-3_9

Paas, F., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. *Educational Psychology Review*, *24*(1), 27–45. https://doi.org/10.1007/s10648-011-9179-2

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology, 8,* Article 422. https://doi.org/10.3389/fpsyg.2017.00422

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

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

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child Development*, *88*(5), 1653–1670.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). San Diego, CA: Academic Press. https://doi.org/10.1016/B978-012109890-2/50043-3

Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory into Practice*, *41*(4), 219–225.

Pintrich, P. R. (2003). A Motivational Science Perspective on the Role of Student Motivation in Learning and Teaching Contexts. *Journal of Educational Psychology*, *95*(4), 667–686. https://doi.org/10.1037/0022-0663.95.4.667

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, *16*(4), 385–407.

Pintrich, P. R., Smith, D. A. F., 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.

Pintrich, P. R., & Zusho, A. (2002). *The development of academic self-regulation: The role of cognitive and motivational factors.* In A. Wigfield & J. S. Eccles (Eds.), *A Vol. in the educational psychology series. Development of achievement motivation* (p. 249–284). Academic Press. https://doi.org/10.1016/B978-012750053-9/50012-7

Ponnock, A., Muenks, K., Morell, M., Seung Yang, J., Gladstone, J. R., & Wigfield, A. (2020). Grit and conscientiousness: Another jangle fallacy. *Journal of Research in Personality*, *89*, 104021. https://doi.org/https://doi.org/10.1016/j.jrp.2020.104021

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*(2), 322–338. https://doi.org/10.1037/a0014996

Puustinen, M., & Pulkkinen, L. (2001). Models of self-regulated learning: A review. *Scandinavian Journal of Educational Research*, *45*(3), 269–286.

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

Rimfeld, K., Kovas, Y., Dale, P. S., & Plomin, R. (2016). True grit and genetics: Predicting academic achievement from personality. *Journal of Personality and Social Psychology*, *111*, 780–789. https://doi.org/10.1037/pspp0000089

Roberts, B. W., Jackson, J. J., Fayard, J. V., Edmonds, G., & Meints, J. (2009). *Conscientiousness.* In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (p. 369–381). The Guilford Press.

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, *2*(4), 313–345.

Roberts, B. W., & Mroczek, D. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science*, *17*(1), 31–35.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin, 132*(1), 1–25. https://doi.org/10.1037/0033-2909.132.1.1

Rottinghaus, P. J., Larson, L. M., & Borgen, F. H. (2003). The relation of self-efficacy and interests: A meta-analysis of 60 samples. *Journal of Vocational Behavior*, *62*(2), 221–236.

Rutherford, T., Buschkuehl, M., Jaeggi, S. M., & Farkas, G. (2018). Links between achievement, executive functions, and self-regulated learning. *Applied Cognitive Psychology*, *32*(6), 763–774.

Salomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology, 76*(4), 647–658. https://doi.org/10.1037/0022-0663.76.4.647

Scheiter, K., Gerjets, P., Vollmann, B., & Catrambone, R. (2009). The impact of learner characteristics on information utilization strategies, cognitive load experienced, and performance in hypermedia learning. *Learning and Instruction*, *19*(5), 387–401. https://doi.org/https://doi.org/10.1016/j.learninstruc.2009.02.004

Schunk, D. H., & Greene, J. A. (2018a). Handbook of Self-Regulation of Learning and Performance. In *Handbook of Self-Regulation of Learning and Performance* (2nd Editio). New York: Routledge. https://doi.org/10.4324/9781315697048

Schunk, D. H., & Greene, J. A. (2018b). Historical, contemporary, and future perspectives on self-regulated learning and performance. In *Educational Psychology Handbook Series. Handbook of self-regulation of learning and performance, 2nd ed.* (pp. 1–15). New York, NY, US: Routledge/Taylor & Francis Group.

Schwab, F., Hennighausen, C., Adler, D. C., & Carolus, A. (2018). Television is still “easy” and print is still “tough”? More than 30 years of research on the amount of invested mental effort. *Frontiers in Psychology*, *9*, 1098. doi: https://doi.org/10.3389/fpsyg.2018.01098

Sevcenko, N., Ninaus, M., Wortha, F., Moeller, K., & Gerjets, P. (2021). Measuring cognitive load using in-game metrics of a serious simulation game. *Frontiers in Psychology*, *12*, 906.

Shim, S., & Ryan, A. (2005). Changes in self-efficacy, challenge avoidance, and intrinsic value in response to grades: The role of achievement goals. *The Journal of Experimental Education*, *73*(4), 333–349.

Sidi, Y., Shpigelman, M., Zalmanov, H., & Ackerman, R. (2017). Understanding metacognitive inferiority on screen by exposing cues for depth of processing. *Learning and Instruction*, *51*, 61–73.

Sitzmann, T., & Ely, K. (2011). A meta-analysis of self-regulated learning in work-related training and educational attainment: What we know and where we need to go. *Psychological Bulletin, 137*(3), 421–442. https://doi.org/10.1037/a0022777

Suchy, Y. (2009). Executive functioning: Overview, assessment, and research issues for non-neuropsychologists. *Annals of Behavioral Medicine*, *37*(2), 106–116.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*(4), 295–312. https://doi.org/https://doi.org/10.1016/0959-4752(94)90003-5

Sweller, J. (2005). *Implications of Cognitive Load Theory for Multimedia Learning.* In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (p. 19–30). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.003

Taub, M., & Azevedo, R. (2019). How does prior knowledge influence eye fixations and sequences of cognitive and metacognitive SRL processes during learning with an intelligent tutoring system? *International Journal of Artificial Intelligence in Education*, *29*(1), 1–28.

Taub, M., Azevedo, R., Bouchet, F., & Khosravifar, B. (2014). Can the use of cognitive and metacognitive self-regulated learning strategies be predicted by learners’ levels of prior knowledge in hypermedia-learning environments? *Computers in Human Behavior*, *39*, 356–367.

The World Bank Group. (2011). 2011). *Learning for all: Investing in people’s knowledge and skills to promote development: Education sector strategy 2020*. Washington, DC: The World Bank. https://doi.org/10.1086/675413

Theobald, M. (2021, January 10). Self-Regulated Learning Training Programs Enhance University Students’ Academic Performance, Self-Regulated Learning Strategies, and Motivation: A Meta-Analysis. https://doi.org/10.31234/osf.io/tf8nk

Thillmann, H., Künsting, J., Wirth, J., & Leutner, D. (2009). Is it merely a question of “what” to prompt or also “when” to prompt? The role of point of presentation time of prompts in self-regulated learning. *Zeitschrift Für Pädagogische Psychologie*, *23*(2), 105–115.

Titz, C., & Karbach, J. (2014). Working memory and executive functions: effects of training on academic achievement. *Psychological Research*, *78*(6), 852–868.

Trautwein, U., Lüdtke, O., Nagy, N., Lenski, A., Niggli, A., & Schnyder, I. (2015). Using individual interest and conscientiousness to predict academic effort: Additive, synergistic, or compensatory effects? *Journal of Personality and Social Psychology, 109*(1), 142–162. https://doi.org/10.1037/pspp0000034

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154.

Tze, V. M. C., Daniels, L. M., & Klassen, R. M. (2016). Evaluating the relationship between boredom and academic outcomes: A meta-analysis. *Educational Psychology Review*, *28*(1), 119–144.

Van Merrienboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, *17*(2), 147–177.

Vedel A., Poropat A.E. (2017) Personality and Academic Performance. In: Zeigler-Hill V., Shackelford T. (eds) *Encyclopedia of Personality and Individual Differences*. Springer, Cham. https://doi.org/10.1007/978-3-319-28099-8_989-1

Vygotsky, L. (1978). Interaction between learning and development. *Readings on the Development of Children*, *23*(3), 34–41.

Whelan, R., & Garavan, H. (2014). When optimism hurts: inflated predictions in psychiatric neuroimaging. *Biological Psychiatry*, *75*(9), 746–748.

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81.

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* (p. 657–700). John Wiley & Sons, Inc.. https://doi.org/10.1002/9781118963418.childpsy316

Winne, P. H. (1995). Self-regulation is ubiquitous but its forms vary with knowledge. *Educational Psychologist*, *30*(4), 223–228.

Winne, P. H. (2001). *Self-regulated learning viewed from models of information processing.* In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (p. 153–189). Lawrence Erlbaum Associates Publishers.

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

Winne, P. H., & Hadwin, A. (2008). The weave of motivation and self-regulated learning. In D. Schunk & B. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 297–314). Mahwah, NJ: Erlbaum.

Winne, P. H., & Perry, N. E. (2000). *Measuring self-regulated learning.* In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (p. 531–566). Academic Press. https://doi.org/10.1016/B978-012109890-2/50045-7

Xie, H., Wang, F., Hao, Y., Chen, J., An, J., Wang, Y., & Liu, H. (2017). The more total cognitive load is reduced by cues, the better retention and transfer of multimedia learning: A meta-analysis and two meta-regression analyses. *PloS One*, *12*(8), e0183884.

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*, *12*(6), 1100–1122.

Yeager, David S, Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., … Hinojosa, C. P. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature*, *573*(7774), 364–369.

Yeager, David Scott, & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist*, *47*(4), 302–314.

Zeidner, M. (2019). Self-regulated learning: Current fissures, challenges, and directions for future research. *High Ability Studies*, *30*(1–2), 255–276. https://doi.org/10.1080/13598139.2019.1584034

Zeidner, M., Boekaerts, M., & Pintrich, P. R. (2000). *Chapter 23 - Self-Regulation: Directions and Challenges for Future Research* (M. Boekaerts, P. R. Pintrich, & M. B. T.-H. of S.-R. Zeidner, Eds.). San Diego: Academic Press. https://doi.org/https://doi.org/10.1016/B978-012109890-2/50052-4

Zelazo, P. D., & Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives*, *6*(4), 354–360.

Zheng, L. (2016). The effectiveness of self-regulated learning scaffolds on academic performance in computer-based learning environments: A meta-analysis. *Asia Pacific Education Review*, *17*(2), 187–202.

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

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology*, *81*(3), 329–339.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In *Handbook of self-regulation.* (pp. 13–39). San Diego, CA, US: Academic Press. https://doi.org/10.1016/B978-012109890-2/50031-7

Table A
Overview of key findings of meta-analyses on SRL

Study	Focus	Population	Outcomes	Sample	Average Effect Size	Moderators / Predictors	Key findings
Broadbent & Poon (2015)	Effect of specific SRL strategies on academic achievement in online or web-based courses	College and university students/ higher education	Academic achievement	12 Studies	Combined: r = .13	Metacognition Time management Effort regulation Critical Thinking Rehearsal Elaboration Organization	SRL strategies correlated with academic achievement: • Metacognition • Time management • Effort regulation • Critical Thinking
Brydges et al. (2015)	SRL in simulation-based training	Medical students, Nurses & nursing students, Postgraduate medical trainees	Learning gains	32 Studies N = 2482		Presence of instructor Test timing (immediate vs retention)	• Unsupervised training vs. no intervention • Supported SRL leads to larger effect sizes then unsupported SRL • Studies designed to support SRL have small beneficial effects over unsupervised studies for immediate posttest and retention tests • Supervised training favors posttest results • Unsupervised training favors retention tests • Specific SRL supports may have benefits and supervision does not consistently improve SRL training outcomes

Table A (continued)

Study	Focus	Population	Outcomes	Sample	Average Effect Size	Moderators / Predictors	Key findings
de Boer et al. (2013)	Effectiveness of SRL interventions	Primary and secondary school students	Students performance • Intervention independent • Self-developed	55 Studies	Overall: g = 0.66	Subject domain Measurement instrument Student characteristics Grade level Implementor Duration of intervention Cooperation Computer use Specific learning strategies	SRL is effective across subject domains (most in writing) Regardless of • students' characteristics (SES) • Grade level (measurement instrument) • The implementor (larger ES if not teacher) The effectiveness of specific learning strategies varies between domains
Dent & Koenka (2016)	Effect of metacognitive and cognitive SRL processes on academic achievement	Elementary and secondary school students	Academic achievement	61 studies 59 studies	Metacognitive processes: r = 0.20/0.27 Cognitive processes: r =0.08/0.11	Specific process Academic subject Grade level Type of SRL measure Type of achievement measure	Correlations between SRL processes are stronger: • For online measures of SRL • In social sciences than in math Metacognitive processes: • Fluctuate over grade levels • Are Highest for standardized tests and weakest for GPA • Composite measures are stronger than single processes Cognitive processes: • Increase with grade level • Highest for GPA and weakest for standardized tests • Stronger for deeper levels of processing

Table A (continued)

Study	Focus	Population	Outcomes	Sample	Average Effect Size	Moderators / Predictors	Key findings
Dignath & Buettner (2008)	Effectiveness of SRL interventions	Primary school Students Secondary school students	Academic performance Strategy use Motivational outcomes	Overall: *N* = 8,619 Primary school: 49 Studies Secondary school: 35 Studies	Overall primary school: 0.68 secondary school: 0.71	Theoretical background of the Intervention Focus of the training instructions Conductor of the study Content domain Group work Length of training	Overall: SRL interventions are more helpful: • in mathematics • for instructions by researchers • for higher number of training sessions Trainings in primary school were more effective for interventions: • based in socio-cognitive theories • Cognitive strategies more than metacognitive reflection in math, but opposite for strategy use Secondary school: • Based on metacognitive theories • Motivation strategies and metacognitive reflection more effective than cognitive strategies
Dignath, Buettner & Langfeldt (2008)	SRL training programs in primary school	Primary school students	Academic performance • Math • Reading & Writing • Other Strategy use • Cognitive or metacognitive • motivational	48 Studies	Overall: d = 0.69 Academic performance: d = 0.62	Content related • cognitive • metacognitive • metacognitive reflection • motivational Context related • Content domain • Duration • Conductor of the study • Age	Trainings were more effective when they • had a social cognitive background • Included more than just cognitive strategies (e.g., planning) • Covered knowledge about strategies And for • Math or reading & writing • Trainings directed by researchers • Younger students

Table A (continued)

Study	Focus	Population	Outcomes	Sample	Average Effect Size	Moderators / Predictors	Key findings
Jansen et al. (2019)	Mediating role of SRL between SRL-Interventions and academic achievement in higher education	University students (higher education)	Academic Achievement SRL activities	142 studies	SRL interventions → achievement: β = 0.18 d = 0.48 SRL interventions → SRL β = 0.22 d = 0.50 SRL → achievement β = 0.22 r = 0.23 SRL interventions → SRL → achievement β = 0.05	Study characteristics • Academic subject • Educational setting • Study design quality • Context Measurement characteristics • SRL activity measured • Measurement instrument SRL • Achievement measure Intervention characteristics • Inclusion of cognitive strategies • Format • Timing • Tailored to the learning context • SRL activity supported	• SRL interventions and activities have a medium impact on achievement • SRL activities only moderate a small part of the effect of SRL interventions on achievement • Interventions have a stronger effect on achievement in humanities than in social sciences • Smaller effect for GPA than course performance • Interventions have a stronger effect on SRL activities when count measures are used (compared to questionnaires) • Solely focusing on resource management is associated with smaller effects of interventions on SRL and SRL on achievement

Table A (continued)

Study	Focus	Population	Outcomes	Sample	Average Effect Size	Moderators / Predictors	Key findings
Richardson et al. (2012)	Nonintellectual correlates of GPA	Students in higher education	Grade point average	217 studies	No average effect • Largest correlation r = 0.59 (self-efficacy)	Predictors (50) • Personality traits • Motivational factors • Self-regulatory learning strategies • Approaches to learning • Contextual influences • Demographic factors	Personality traits • Procrastination • Conscientiousness • Need for cognition • Emotional intelligence Motivational factors • Locus of control • Intrinsic motivation • Goal orientation Self-regulatory learning strategies • Elaboration • Critical thinking • Use of meta cognition • Help seeking • Time/study management • Concentration • Test anxiety Approaches to learning • Strategic • Deep • Surface Psychosocial contextual influences • Goal commitment • General stress or stress relating to university work

Table A (continued)

Study	Focus	Population	Outcomes	Sample	Average Effect Size	Moderators / Predictors	Key findings
Sitzman & Ely (2011)	SRL trainings and work-related knowledge and skills	Trainees • University students • Employees • Military personnel	Training transfer	369 studies N = 90,380	No average effect: Corrected r [-.30; .83]	Predictors: Regulatory agent Regulatory mechanisms Regulatory appraisals Moderators: Study population Research design Publication year and type	Strong correlations between different self-regulatory constructs Strongest predictors of learning: • Goal level • Self-efficacy • Effort • Persistence • Attention
Theobald (under review)	SRL interventions in higher education	University students (incl. postgraduates)	Academic performance Cognitive strategies Metacognitive strategies Resource management strategies Motivational outcomes	32 Studies N = 4106	Overall: g = 0.36	Feedback Cooperative learning Learning protocols Age Prior achievement	SRL interventions foster academic performance • Lower if GPA is used as outcome measure SRL interventions foster strategy use • Metacognitive and resource management strategies more than cognitive Effects on cognitive and metacognitive strategy use are stronger in cooperative learning settings Resource management strategy use is stronger with teacher feedback and learning protocols • Further moderated by age and prior knowledge SRL interventions foster motivation • intrinsic motivation & interest • self-efficacy

Additional analyses for Study III

Research question. Do emotion profiles (i.e., positive, neutral, negative) significantly differ in their personality (agreeableness, conscientiousness, extraversion, intellect, and neuroticism)?

Results. To test if the emotion profiles from the three-profile solution (i.e., positive, neutral, negative profile) significantly differed in their personality (agreeableness, conscientiousness, extraversion, intellect, and neuroticism measured with the short form of the International Personality Item Pool, Donnellan et al., 2006) a multivariate analysis of variance comparing mean levels of the five personality facets between emotion profiles was conducted. Results showed that emotion profiles significantly differed in their personality (Wilks's λ (10, 162) = 0.100, p < .001, $\eta_p{}^{2}$ = 0.09). Follow-up analyses split by personality facets demonstrated that emotion profiles significantly differed in neuroticism (F(2, 162) = 8.52, p < .001, η^2 = .10) but on none of the other facets (Fs(2, 162) ≤ 2.76, ps ≥ .067, $\eta_p{}^2$ ≤ .03). Lastly, post-hoc pairwise t-test with Bonferroni corrected p-values showed that the negative emotion profile was characterized by significantly higher levels of neuroticism (M = 3.16, SD = 0.73) then the neutral (M = 2.72, SD = 0.82, p = .025) and positive profile (M = 2.44, SD = 0.87, p = .001). No significant differences were found between the neutral and the positive emotion profile (p = .195).