Environmental and Learner-Related Determinants of Knowledge Acquisition:

Meta-Analytic and Longitudinal Analyses

Dissertation

submitted in fulfillment of the requirements for the degree of

doctor rerum naturalium (Dr. rer. nat.) in Psychology

at the University of Trier

by

Bianca A. Simonsmeier

Supervised by

Prof. Dr. Michael Schneider

Prof. Dr. Franzis Preckel

Submitted on December 19th 2017

Place of Dissertation: Trier, Germany

Acknowledgements

After intensive 33 months of work on my dissertation, I would like to reflect on the people who have supported and helped me so much throughout the whole period.

Working with great people is one of the most important things for great work. I am more than thankful to be part of a wonderful lab providing me with the best working environment I can imagine. I hereby want to express my sincere gratitude to my advisor Prof. Dr. Michael Schneider for his continuous support, limitless patience, extensive knowledge, and maddening attention to detail. I would like to thank Prof. Dr. Franzis Preckel for supervising my work, her valuable input, and inspiring and motivating me with her achievements, serving as a great role model. I also want to thank my co-workers and friends Susanne Bücker, Maja Flaig, and Sevim Nuraydin for their inputs and emotional support throughout the last three years.

Another nourish for my work came from my family, who I deeply thank for their extensive support. I want to thank my partner Ruven for showing me how to relax and how to relieve my mind, even in the toughest situations. I want to thank my grandparents Klaus and Helene for providing me with the feeling that I can achieve anything in my life I thrive for. I want to thank my grandparents Gustav and Inge for supporting my studies financially and giving me the opportunity to finally find my passion in research and educational psychology. I want to thank my mother for always having an open ear and being there for me, no matter what. I finally want to thank my father for encouraging me to give ones best and never give up, regardless of any ups and downs.

Abstract

The search for relevant determinants of knowledge acquisition has a long tradition in educational research, with systematic analyses having started over a century ago. To date, a variety of relevant environmental and learner-related characteristics have been identified, providing a wide body of empirical evidence. However, there are still some gaps in the literature, which are highlighted in the current dissertation. The dissertation includes two meta-analyses summarizing the evidence on the effectiveness of electrical brain stimulation and the effects of prior knowledge on later learning outcomes and one empirical study employing latent profile transition analysis to investigate the changes in conceptual knowledge over time. The results from the three studies demonstrate how learning outcomes can be advanced by input from the environment and that they are highly related to the students' level of prior knowledge. It is concluded that the effects of environmental and learner-related variables impact both the biological and cognitive processes underlying knowledge acquisition. Based on the findings from the three studies, methodological and practical implications are provided, followed by an outline of four recommendations for future research on knowledge acquisition.

Table of Contents

1. Introduction

Knowledge is a precious commodity we aim to amass, preserve, and disseminate (Hulbert & Anderson, 2008). Researchers and educators across various disciplines have been devoted to scientifically studying and deeply understanding the processes and the mechanisms underlying knowledge acquisition, and how these can be systematically supported (e.g., de Corte, 2010). It is widely assumed that knowledge acquisition occurs through intra-learner processes (e.g., George A. Miller, 1956; Sweller, van Merriënboer, & Paas, 1998), and is facilitated by environmental factors such as information provided in the form of teaching and instruction (Gagné, 2004; Stern, 2017; Sweller et al., 1998; Ullén, Hambrick, & Mosing, 2016) as well as learner-related factors such as prior knowledge or general abilities (e.g., Baddeley, Eysenck, & Anderson, 2009; Biggs, 2003; Gagné, 2004; Stern, 2017; Ullén et al., 2016). The current dissertation adopted this view and investigated the effects of environmental and learner-related determinants on knowledge acquisition. There are three main sections in this dissertation: the introduction and theoretical background (Chapters 1 to 4), the presentation of the three studies as the core of this dissertation (Chapters 5 to 7), and the general discussion (Chapter 8) and conclusion (Chapter 9). More precisely, the findings from the single studies provide comprehensive insights regarding the effects of transcranial electrical stimulation on school relevant learning outcomes (Study 1), the relevance of prior knowledge on learning outcomes and environmental and learner-related moderating variables of this relationship (Study 2), and transition pathways of the acquisition of conceptual knowledge over time considering different qualities of knowledge (Study 3).

2. Knowledge

The concept of knowledge originally stems from early philosophy and is now used in multiple domains and disciplines. One of the oldest and most widely used definitions of knowledge originated from Plato's Theaetetus stating that knowledge is "justified true belief". While this definition is not without problems, it points to the fact that knowledge is a mental state. Many modern-day psychological theorists subscribe to this view and use the words knowledge and mental representation interchangeably (e.g., Markman & Dietrich, 2000; Spelke & Kinzler, 2007). In line with this, knowledge is mentally represented information stored in memory (e.g., Anderson, 1983; Weinert, 1999), which is ultimately instantiated in neuronal connections in the brain (Brod, Werkle-Bergner, & Shing, 2013; Poldrack et al., 2011). As such, knowledge is a person-related variable describing information stored in memory (Yates & Chandler, 1991), that can vary substantially across individuals.

The initial status of knowledge, also referred to as prior knowledge, is relevant for further knowledge acquisition and resulting broader academic outcomes (e.g., Stanovich, 1986; Watts, Duncan, Siegler, & Davis-Kean, 2014). Prior knowledge is defined as the knowledge available in a person's long-term memory before learning something new (cf. Dochy & Alexander, 1995). It is useful to distinguish knowledge from academic achievement, which has been commonly investigated in educational studies. Academic achievement is defined as the extent to which a learner has achieved short or long-term educational goals that were the focus of activities in instructional environments, for example in primary and secondary school, or college and university (Steinmayr, Meißner, Weidinger, & Wirthwein, 2014). Academic achievement can be assessed either by general indicators such as domain-specific knowledge measures, more curricular-based criteria (e.g., grades or performance on an achievement test), or by means of cumulative indicators (e.g., degrees or educational certificates) (Steinmayr et al., 2014). Thereby, knowledge denotes a component of broader cognitive learning outcomes such as achievement (Steinmayr et al., 2014), literacy (OECD,

2016), or expert performance (Gobet, 1998). Achievement, competence, or expertise is based on the accumulation of knowledge over time (Brod et al., 2013; Yates & Chandler, 1991).

2.1 Characteristics of Knowledge

Conceptually, knowledge can be distinguished between various types of knowledge and different aspects of the quality of knowledge (de Jong & Ferguson-Hessler, 1996; Farnham-Diggory, 1994). Types of knowledge may include declarative or conceptual knowledge (e.g., facts, concepts, principles; Goldwater & Schalk, 2016; Michael Schneider & Stern, 2010), procedural knowledge (i.e., knowledge about actions which are important for problem solving; Anderson et al., 2004; Michael Schneider & Stern, 2010), situational knowledge (i.e., knowledge about typical problem situations; de Jong & Ferguson-Hessler, 1996), or logical knowledge (i.e., a mental model of what is connected to what and what leads to what; Farnham-Diggory, 1994). Qualities of knowledge can include the hierarchical organization of knowledge (superficial vs. deeply embedded; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976; Rumelhart, 1980), the inner structure of knowledge (isolated knowledge elements vs. well structured, interlinked knowledge; diSessa, 2008; diSessa, Gillespie, & Esterly, 2004; Michael Schneider & Grabner, 2012), the level of automation (i.e., declarative vs. compiled; Anderson et al., 2004; Sweller et al., 1998), the level of abstraction (colloquial vs. formal; Krems, 1994), the generalizability of knowledge (general vs. domain-specific; Carey & Spelke, 1994; Wellman & Gelman, 1992), the explicitness of knowledge (explicit vs. implicit; Dienes & Perner, 1999), or its correctness (e.g., incorrect vs. partially correct vs. scientifically correct, Shtulman & Valcarcel, 2012; Smith III., diSessa, & Roschelle, 1993; Vosniadou & Brewer, 1992). Although different types or qualities of knowledge cannot always be empirically separated (e.g., Rittle-Johnson & Schneider, 2015), a theoretical distinction can be helpful in terms of achieving a better understanding of knowledge acquisition over time and possible moderating variables of these developments.

2.2 Storage of Knowledge: Connectivity Patterns and Cognitive Architectures

A wide body of research has been devoted to gain a deep understanding on how knowledge is stored. Research mainly stems from two different disciplines, namely neuroscience and cognitive science. From a biological perspective, knowledge is stored in form of modified synaptic connectivity. It is assumed that a specific connectivity pattern between synapses of a set of neurons is linked to a mental representation. The connection between the connectivity pattern and the mental representation is also called an engram. It is thought that information is stored in the very same cell assemblies that are active during online information processing of mental activities. Therefore, input of new information substantiates a particular activation pattern, as it resonates with the established connections (Rösler & Heil, 2003).

From a cognitive perspective, knowledge is stored in long-term memory. There are various models describing the cognitive architectures of knowledge storage (e.g., Anderson et al., 2004; Bartlett, 1932; Newell, 1990). They share the basic assumption that knowledge is stored in long-term-memory in the form of elements which may either be related or not, commonly referred to as associative networks. This implies that knowledge structures are virtually unlimited in their size, complexity, and sophistication and that they are built over time (Bartlett, 1932; Chi, Glaser, & Rees, 1982; Rumelhart & Orthony, 1977).

Understanding the structures that constitute human cognitive architectures can serve as a foundation for instructional approaches and for achieving a deep understanding of the processes of knowledge acquisition in educational settings (Kirschner, Sweller, & Clark, 2006; H. G. Schmidt, Loyens, van Gog, & Paas, 2007).

2.3 Acquisition of Knowledge

Knowledge acquisition “is conceived in terms of storage of information in memory as a consequence of any experience the individual might have had” (Medin & Ross, 1992, p. 55). This definition has two major implications. First, knowledge acquisition means changes to

information stored in long-term memory. Second, various (intra-learner) processes of change take place during knowledge acquisition (i.e., acquisition, modification, enrichment of knowledge), and these processes may be determined by the characteristics of the learner and their environment. The following section will reflect this logic by first examining underlying intra-learner processes of knowledge acquisition, followed by an overview of relevant learner characteristics and influences from the learners' environment.

2.3.1 Intra-Learner Processes

A wide body of research has demonstrated the relevance of intra-learner processes for knowledge acquisition on both neurophysiological and cognitive levels. These processes and their interaction will be briefly presented in the paragraphs that follow.

From a neurophysiological perspective, knowledge acquisition is the reorganization of brain structures, which is also known as neuroplasticity. This change may be described on different levels, ranging from synaptic modification, to how populations of neurons cooperate, to how multiple brain areas interact (Newman & Polk, 2008). There are different mechanisms underlying neuroplasticity, namely synaptic plasticity, nonsynaptic plasticity, and metaplasticity (Abraham, 2008; Daoudal & Debanne, 2003). Synaptic plasticity describes the process by which synapses are modified in their structure and function due to incoming stimuli or cues from the environment (Gaiarsa, Caillard, & Ben-Ari, 2002). The classical models underlying synaptic plasticity are long-term potentiation (LTP), denoting an increase in synaptic strength, and long-term depression (LTD), denoting a decrease in synaptic strength (Bear & Malenka, 1994). Nonsynaptic plasticity describes the modification of a single neurons' intrinsic electrical properties (Daoudal & Debanne, 2003), often also referred to as intrinsic plasticity. The models underlying nonsynaptic plasticity are E-S potentiation and E-S depression, denoting an excitatory postsynaptic potential, which can either be decreased or increased (Bliss & Lømo, 1973; Daoudal, Hanada, & Debanne, 2002). Synaptic plasticity and nonsynaptic plasticity are both bidirectional and functionally synergetic

(Daoudal & Debanne, 2003). Metaplasticity describes activity-dependent changes in neural functions that modulate subsequent synaptic plasticity (Abraham & Bear, 1996). It refers to neuronal changes that occur at one point in time which is commonly called priming activity, regulating processes of synaptic plasticity (Abraham, 2008).

A major contribution of cognitive science to educational theory and research is to provide models of how learners process information to build knowledge in their memory (e.g., Mayer, 2001; Mayer & Moreno, 2003; George A. Miller, 1956; Paivio, 1991; Sweller et al., 1998). The models outline the central role of short-term sensory storage, working memory, and long-term memory. Sensory storage and working memory enable learners to manage limited amounts of incoming information during initial processing, whereas long-term memory serves as a permanent storage for knowledge. The short-term sensory storage screens incoming stimuli and transfers only the most relevant stimuli to working memory. Basic cognitive processes relevant for this step are attention and perception. Working memory serves as a multi-component temporary storage, where new information is assigned meaning, linked to other information, and essential mental operations, such as inferences, are performed. Basic cognitive processes executed in working memory are selective attention, selective processing, and elaboration of new information. Long-term memory serves as a permanent repository for different types and qualities of knowledge (e.g., procedural knowledge or declarative knowledge and fragmented or complete knowledge). Working memory and long-term memory are connected by encoding and retrieval processes. Encoding refers to the strategies that describe how information from the temporary storage in working memory is transferred into long-term memory (i.e., knowledge acquisition). Retrieval describes the process of searching memory and accessing stored information (i.e., prior knowledge) for active processing in working memory. Both encoding and retrieval facilitate knowledge acquisition when information in long-term memory is easily accessible.

2.3.2 Environmental Determinants

Knowledge acquisition is highly determined by input from the environment (e.g., Mayer, 2001; Stern, 2017), raising the question of how the learning environment can enhance knowledge acquisition and its underlying biological and cognitive processes. The search for effective instruction as an advancement for knowledge acquisition has attracted research for over a century (Alexander & Mayer, 2011). However, the effectiveness of instruction may be limited by biological factors (Blakemore & Choudhury, 2006; Daoudal & Debanne, 2003; Mua, Liac, Yaob, & Zhuoc, 1999) or cognitive factors (e.g., Mayer & Moreno, 2003; George A. Miller, 1956; Sweller et al., 1998; Tyng, Amin, Saad, & Malik, 2017). One method targeting the enhancement of biological processes underlying knowledge acquisition is electrical stimulation of the brain (e.g., Moreno-Duarte et al., 2014). As knowledge can also be seen as activity patterns in the brain that change due to learning (Daoudal & Debanne, 2003; Gaiarsa et al., 2002), manipulation of these activity patterns may possibly be beneficial for the enhancement of knowledge acquisition. Cognitive processes underlying knowledge acquisition may be targeted by instructional processes. It is most beneficial when instruction accounts for the limited cognitive capacity of working memory in humans, as proposed by cognitive load theory (e.g., Sweller et al., 1998). Two examples for environmental determinants of knowledge acquisition will be more closely discussed in the following paragraphs, respectively.

A method targeting the biological processes of neuroplasticity is transcranial electrical stimulation (tES; Moreno-Duarte et al., 2014), where the brain is non-invasively stimulated with a weak and painless electrical current. The current is delivered through electrodes on the scalp and is assumed to modulate the excitability, long-term synaptic plasticity, or oscillatory activity of the brain's underlying cortical regions (Moreno-Duarte et al., 2014). The five most common types of tES are transcranial direct current stimulation (tDCS), high-definition transcranial direct current stimulation (HDtDCS), transcranial alternating current stimulation

(tACS), transcranial pulsed current stimulation (tPCS), and transcranial random noise stimulation (tRNS) (Nitsche & Paulus, 2000). These procedures differ both in their intensity (ranging from -500 μA to 2 mA) as well as in the type of current applied (e.g., constant current, bidirectional, biphasic current, or alternating current). The typical duration of electrical stimulation ranges from 2 to 20 minutes. The effects of tES are usually investigated by comparing a tES group with a sham control group, in which everything is done similarly, including the placement of the electrodes, but no actual electrical stimulation is given for most of the session (for more details, see Dissanayaka, Zoghi, Farrell, Egan, & Jaberzadeh, 2017). During the last few years, several meta-analyses have investigated the effects of tES on neurophysiological and cognitive variables (Horvath, Forte, & Carter, 2015a, 2015b; Price, McAdams, Grossman, & Hamilton, 2015), but the results of the included single studies as well as of the overall results found in the meta-analyses are heterogeneous. Furthermore, none of the previous meta-analyses have investigated the effects of tES on knowledge acquisition for school-relevant tasks. However, this seems promising, as it is assumed that tES enhances neuronal correlates of learning (e.g., Stagg et al., 2009; Stagg & Nitsche, 2011). It is therefore of scientific interest to more deeply investigating the effects of tES, to determine whether advancement of knowledge acquisition using tES is possible.

For the acquisition and reorganization of information in long-term memory, information processing through working memory is a key factor (e.g., Stern, 2017). However, working memory is limited in capacity, potentially restricting the transfer of information from working memory to long-term memory (George A. Miller, 1956). The cognitive load theory (CLT; Sweller et al., 1998) addresses the cognitive processes underlying knowledge acquisition. It proposes that instruction designed to reduce working memory load facilitates learning. There is a wide body of evidence indicating that the reduction of cognitive load is effective, as demonstrated in various single studies and meta-analyses. For example, the reduction of cognitive load has been demonstrated to be effective in terms of error prevention training

(Hutchins, Wickens, Carolan, & Cumming, 2013), the use of instructional animations (Höffler & Leutner, 2007), or an increase of spatial or temporal contiguity of related elements (Ginns, 2006). The positive effects of addressing cognitive load in instruction were consistently found across different domains and specific instructional methods, indicating the generality of the effectiveness when considering cognitive processes for knowledge acquisition.

2.3.3 Learner-Related Determinants

One of the key components of knowledge acquisition is prior knowledge, both from a biological and cognitive perspective (e.g., Brod et al., 2013; Newman & Polk, 2008; Sweller et al., 1998). Neuroscientists argue that prior knowledge directs information processing in a top-down-manner. Prior knowledge as a higher-order representation can affect lower-order processes, for example the recognition of newly presented material (e.g., Newman & Polk, 2008). Further, recent findings from neuroscience point to a prominent role of the medial prefrontal cortex and the hippocampus in the emergence of prior knowledge, and in its application for successful memory encoding, consolidation, and retrieval (Maguire, Frith, & Morris, 1999; van Kesteren, Rijpkema, Ruiter, & Fernández, 2010; van Kesteren, Rijpkema, Ruiter, Morris, & Fernández, 2014; Wang & Morris, 2010). Prior knowledge is also highlighted as a key variable in cognitive theories and models (e.g., Chase & Ericsson, 1982; Stern, 2017; Sweller et al., 1998). As working memory and long-term memory are directly connected, prior knowledge stored in long-term memory affects students' selective attention, selective processing, encoding, retrieval, and storage, which becomes evident in the (re-)structuring and integration of knowledge in long-term memory. From this perspective, learning and thinking processes are never knowledge free, which highlights the dynamic role of knowledge in human information processing.

A remarkable amount of research has considered prior knowledge, its contribution to knowledge acquisition and academic performance, and the underlying cognitive principles. Prior knowledge positively influences knowledge acquisition (e.g., Stanovich, 1986; Yates &

Chandler, 1991). When comparing individuals on knowledge levels, a minor difference at one time point will likely become a large difference on a subsequent assessment. This cumulative impact of prior knowledge is an example of the Matthew effect (i.e., an increase in individual differences such that the rich get richer and the poor get poorer) and has been empirically supported across many learner sub-groups and various domains (Duff, Tomblin, & Catts, 2015; Sharp, Sinatra, & Reynolds, 2008; Stanovich, 1986; Walberg & Tsai, 1983; Watts et al., 2014). Research has suggested several psychological mechanisms to explain the positive effects of prior knowledge on acquiring new knowledge, mainly focusing on enhanced cognitive processes due to prior knowledge (Craik & Bialystok, 2006). For example, previous research has demonstrated positive effects of prior knowledge on processing speed (e.g., Chase & Simon, 1973), information encoding (Crooks & Alibali, 2013; DeWitt, Knight, & Hicks, 2012), retrieval of information (e.g., Chi & Ceci, 1987; Jensen & Hansen, 1995; Recht & Leslie, 1988; W. Schneider, Körkel, & Weinert, 1989), and knowledge integration (e.g., Hambrick & Engle, 2002; Murphy, 2002). Considering different qualities of knowledge, prior knowledge can also inhibit the acquisition of new knowledge (cf. Dochy, Segers, & Buehl, 1999; W. Schneider, 1993; Stern, 2001; Vosniadou & Brewer, 1992), potentially limiting the amount of change for less knowledgeable as compared to more knowledgeable students. It is thought that misconceptions lead to interferences as new information is more likely to be interpreted in terms of such initial misconceptions (e.g., Guzzetti, Snyder, Glass, & Gamas, 1993; van Loon, de Bruin, van Gog, & van Merriënboer, 2013; Vosniadou, 1994; Vosniadou & Skopeliti, 2017). Alternately, more knowledgeable students may experience less interference in knowledge acquisition as they may incorporate less misconceptions. However, much less is empirically known about how misconception at one point relate to subsequent knowledge at a later points in time (e.g., Michael Schneider & Hardy, 2013).

On the other side, the gap between individuals with more or less knowledge may close over time as described by the compensation effect (e.g., Baumert, Nagy, & Lehmann, 2012;

Kalyuga, Ayres, Chandler, & Sweller, 2003; Schroeders, Schipolowski, Zettler, Golle, & Wilhelm, 2016). Compensation effects have been empirically supported in different academic settings (e.g., Baumert et al., 2012; Huang, Moon, & Boren, 2014; Kalyuga, 2008), with various potential mechanisms for this effect having been proposed. Compensatory effects may likely be found when students receive regular school education, which mainly focuses on the promotion of less knowledgeable students, resulting in a lack of support for more knowledgeable students in academic contexts (e.g., Schroeders et al., 2016; Snow, 1989). More specifically, instructional design may have a positive impact on closing the gap between students differing in prior knowledge. This can be achieved by implementing instructional design principles based on cognitive load theory (Sweller et al., 1998) as demonstrated in various studies investigating expertise reversal effects (Kalyuga et al., 2003). It is assumed that instructional guidance is effective in enhancing knowledge acquisition in less knowledgeable learners, as it reduces cognitive load for this particular learner sub-group. However, it is ineffective for knowledgeable learners as it likely increases cognitive load for those students, which hinders knowledge acquisition (Kalyuga et al., 2003).

Summarizing, a wide body of research demonstrates the relevance of prior knowledge for later knowledge acquisition. Specifically, research has found a positive association between prior knowledge and knowledge acquisition (i.e., Matthew effect) as well as a negative association (i.e., compensatory effect, misconception effect), leaving it open whether the overall net effect of prior knowledge on learning outcomes is positive, negative, or close to zero if the positive and negative effects cancel each other out. Further, empirical insights of the Mathew effect mainly stem from studies quantifying the amount of knowledge, leaving out whether differences in the quality affect knowledge acquisition over time in the same manner.

3. Methodological Approaches

Statistical analyses targeting the development of knowledge structures and the aggregation of existing scientific evidence have more recently gained wide attention. A method that allows the quantitative examination of pathways of knowledge acquisition is latent transition analysis (Hickendorff, Edelsbrunner, McMullen, Schneider, & Trezise, in press). The first section of this chapter will briefly introduce this method and provide possible areas of application. Further, meta-analytic synthesis of data (Borenstein, Hedges, Higgins, & Rothenstein, 2009; F. L. Schmidt & Hunter, 2015) has gained considerable interest, as it allows drawing an overall picture of scientific evidence on a specific topic by integrating research findings across single studies. The second section of this chapter targets the principles, advantages, and limitations of meta-analyses.

3.1 Latent Class, Latent Profile, and Latent Transition Analyses

In traditional statistical approaches, it is assumed that the relationship between variables can be similarly applied to all learners. This implies that individual differences across students are homogeneous (Bergman & Magnusson, 1997; Collins & Lanza, 2010). If qualitative distinctions between subgroups exist within a population, consequently they are not accurately represented by this general model (Hickendorff et al., in press). However, individual differences are regularly expected in educational contexts. In these cases, most often cut-off values are used for statistical analyses. For example, differences in prior knowledge across learners are commonly operationalized by dividing the learners into high and low prior knowledge groups. The most prevalent technique is median split (e.g., Amadieu, Tricot, & Marine, 2009; Fyfe, Rittle-Johnson, & DeCaro, 2012; Kalyuga, 2008; U.-M. Krause, Stark, & Mandl, 2009; Kyun & Lee, 2009; Salmeron, Kintsch, & Canas, 2006). Variations between the groups are then investigated in terms of differences in learning outcomes. There a several disadvantages of cut-off points. They are likely to be arbitrary and

can result in a distorted picture of the relationships between variables (e.g., Altman & Royston, 2006; Irwin & McClelland, 2003). An alternative to analyzing heterogeneous populations is the application of latent class analysis (for categorical variables) or latent profile analysis (for continuous variables). These methods are commonly used to identify learner sub-groups that differ on specific variables. The differences can either be quantitative (for example the amount of correct prior knowledge) or qualitative in nature (for example the quality of knowledge such as misconceptions). In latent transition models, pathways between latent classes or profiles of knowledge over time are investigated by adding a transitioning component.

There are several steps for performing latent class, latent profile, and latent transition analysis (for a comprehensive overview, see Hickendorff et al., in press). Depending on the measurement level of the observed variable and the number of measurement points, different latent class and latent profile models exist (for example, a latent class model for analyzing categorical indicators in a cross-sectional design or latent profile transition models for analyzing continuous indicators in a longitudinal design). After deciding for a research question and an appropriate model, the number of latent classes or profiles can be determined based on existing measurement data. If multiple measurement points were assessed, transition analyses may be performed. The classes as well as the transition paths between latent classes or profiles of knowledge can be further predicted by external variables, for example intelligence or socio-economic-status. Therefore, latent transition analyses can be utilized for a broad range of research questions that focus on individual differences and changes over time. They can further be applied to quantitative as well as qualitative data, or a mix of both.

3.2 Meta-Analyses

Meta-analysis describes a quantitative synthesis that pools or aggregates findings from multiple single studies using statistical techniques (Borenstein et al., 2009; F. L. Schmidt &

Hunter, 2015). At best, the single studies included in a meta-analysis have been explicitly and rigorously identified by systematic review techniques (J. P. T. Higgins & Green, 2008; Petticrew & Roberts, 2005). There are various statistical models and approaches underlying meta-analyses which are discussed in the next paragraphs.

Meta-analyses may differ in their underlying statistical models (Cheung, 2015), more specifically (1) fixed effects versus random-effects models, (2) models for independent versus dependent effect sizes, and (3) models with or without structural models on the average effect sizes. First, meta-analyses are commonly employed to estimate a single summary estimate of a hypothetical true underlying effect that each study determined. The identification of a single effect size in a population is performed by means of a fixed-effect model. A more complex, and likely more-realistic meta-analytic strategy, is the application of a random-effects model. This analysis allows the aggregation of evidence for different underlying effects from single studies that vary as a function one or more moderator variables (Carter, Schönbrodt, Gervais, & Hilgard, 2017; Cooper, Hedges, & Valentine, 2009; Raudenbush, 2009). Second, effect sizes may be statistically dependent, and thus violate a central assumption of classical meta-analytic models. For example, this may be the case when a study reports several effect sizes for various dependent measures or measurement points. In these cases, structural equation modeling (e.g., Cheung, 2015) or robust variance estimation (RVE; Hedges, Tipton, & Johnson, 2010; Tanner-Smith & Tipton, 2014; Tanner-Smith, Tipton, & Polanin, 2016) may be applied. Third, researchers may be interested in determining the impact of mediators or moderators on effect sizes. This may be performed by means of meta-regression analyses (e.g., Baker, White, Cappelleri, Kluger, & Coleman, 2009; Hedges et al., 2010; Morton, Adams, Suttorp, & Shekelle, 2004).

Conducting a meta-analysis involves several steps, namely the formulation of a research question (definition of the research goal, the central keywords, inclusion and exclusion criteria, and potential moderators), the performance of a standardized literature search, the

coding of studies (based on specific coding rules), and the meta-analytic procedure (calculation and transformation of effect sizes, aggregation of effect sizes, analysis of moderator effects). How the specific steps are performed can greatly differ across studies. For example, some meta-analyses include both published and unpublished research, whereas other studies include only published research. Further, in some meta-analyses the effect sizes are computed from raw data reported in the single studies, whereas others use the effect sizes reported in the studies.

The results from meta-analyses may systematically over- or underestimate the true effect, which is referred to as bias (Carter et al., 2017; F. L. Schmidt & Hunter, 2015). Meta-analytic bias may be caused by study artefacts or general characteristics of single studies included in the meta-analysis. Common biases due to study artifacts are caused by sampling error, measurement error, dichotomization, range variation, attrition artefacts, deviation from perfect construct validity, reporting or transcriptional error, and variance due to extraneous factors that affect the relationship. More general biases are availability bias, source bias, and publication bias (for an overview on the different biases, see F. L. Schmidt & Hunter, 2015). To facilitate better interpretation of the results of a meta-analysis, the magnitude of bias can be determined (e.g., Duval & Tweedie, 2000; Egger, Smith, Schneider, & Minder, 1997; Henmi & Copas, 2010). Further, the biases can be addressed statistically by applying corrections to the effect sizes obtained from single studies (e.g., F. L. Schmidt & Hunter, 2015).

In sum, meta-analyses allow for the analysis of patterns of relationships and moderator analyses by combining (inconsistent) evidence from single studies into an overall effect size. Results from meta-analyses may serve as the foundation for theory validation, theory development, and policy formulation by accumulating the knowledge gained from the existing scientific evidence within a specific field and by providing new insights which are not inferable from individual studies (Hoffert, 1997; Hunter & Schmidt, 1996; F. L. Schmidt

& Hunter, 2015). Problems with meta-analyses can occur when aggregating data from different designs, treatments, or constructs, raising the question whether the results from differing studies are still comparable and if the obtained results could be misleading. As in single studies, results from meta-analyses may be biased. However, meta-analyses have the advantages of a more accurate estimation of the overall effect and the analysis of mediating and moderating variables.

4. The Current Dissertation

Knowledge acquisition highly depends on intra-learner processes, which can be described and analyzed on a biological and cognitive level. These intra-learner processes are influenced by external factors from the environment, such as systematically applied advancements of knowledge acquisition (i.e., tES, instruction), and learner-related factors, such as the individuals' prior knowledge stored in long-term memory. Both environmental as well as learner-related characteristics can significantly contribute to learning outcomes such as knowledge or academic achievement. This dissertation aimed to investigate the effects of environmental characteristics and prior knowledge, as a learner characteristic, on knowledge acquisition by means of two meta-analyses and one latent transition analysis.

The first study provides insights in regards of whether tES affects learning outcomes based on a meta-analytic aggregation of existing evidence. It was assumed that tES, as an environmental factor, influences biological substrates of knowledge acquisition, which are associated with learning outcomes (Stagg et al., 2009; Stagg & Nitsche, 2011). Therefore, the effects of tES were expected to be evident in school-relevant learning outcomes. Overall, the study provides the first insights on the effects of tES on knowledge acquisition and potential third variables influencing the effects.

The second study investigated the relationship between prior knowledge and learning outcomes in a meta-analysis. It was assumed that prior knowledge has a significant influence

on the cognitive process of information processing, supporting the acquisition and reformation of knowledge (Anderson, 1983; Stern, 2017; Sweller et al., 1998). This may lead to the accumulated advantage of students with high prior knowledge (Stanovich, 1986; Yates & Chandler, 1991). However, differences in prior knowledge can also be reduced over time (Baumert et al., 2012; Kalyuga et al., 2003), especially when instruction accounts for the demands of less knowledgeable students (Sweller et al., 1998). The second study investigated whether the overall net effect of prior knowledge on learning outcomes is positive, negative or close to zero, and whether the effect was moderated by characteristics of the knowledge, the learner, or the environment.

Study 3 investigated the relationship between prior knowledge and learning outcomes longitudinally using latent transition analysis. Despite the advantage of the amount of prior knowledge for knowledge acquisition, much less is known about pathways of knowledge acquisition starting with incorrect or partly incorrect knowledge as is manifested in misconceptions. As an essential part of prior knowledge, misconceptions are very common (Masson, Potvin, Riopel, & Brault-Foisy, 2014). In the current study, conceptual knowledge was assessed over the course of two semesters, allowing for the investigation of individual pathways of knowledge acquisition in students differing in the quality of knowledge structures.

5. Study 1: Electrical Brain Stimulation (tES) Improves Learning More Than Performance: A Meta-Analysis

5.1 Abstract

Researchers have recently started evaluating whether stimulating the brain noninvasively with a weak and painless electrical current (*transcranial Electrical Stimulation*, tES) enhances physiological and cognitive processes. Some studies found that tES has weak but positive effects on brain physiology, cognition, or assessment performance, which have attracted massive public interest. We present the first meta-analytic test of the hypothesis that tES in a learning phase is more effective than tES in an assessment phase. The meta-analysis included 246 effect sizes from studies on language or mathematical competence. The effect of tES was stronger when stimulation was administered during a learning phase (d = 0.712) as compared to stimulation administered during test performance (d = 0.207). The overall effect was stimulation-dosage specific and, as found in a previous meta-analysis, significant only for anodal stimulation and not for cathodal. The results provide evidence for the modulation of long-term synaptic plasticity by tES in the context of practically relevant learning tasks and highlight the need for more systematic evaluations of tES in educational settings.

5.2 Theoretical Background

The search for more effective learning techniques has been a quest of humanity since time immemorial. Recently, researchers have started evaluating whether stimulating the brain non-invasively with a weak and painless electrical current (*transcranial Electrical Stimulation*, tES) enhances learning and performance. In tES, typically two electrodes are attached to the scalp via saline-soaked sponges. The applied current is usually not higher than 2 mA. In a stimulation session, the current is typically applied from 2 minutes to 20 minutes. The most common form of tES is transcranial Direct Current Stimulation (tDCS), where

anodal and cathodal stimulation can be distinguished. The effects of tES are usually investigated by comparing a tES group with a SHAM control group, in which everything is done the same, including the placement of the electrodes on the skull, but no actual electrical stimulation is given for most of the session. To many, the attractiveness of tES research with healthy participants lies in the promise of achieving performance improvements without investing more time and effort. tES has attracted massive public interest as demonstrated by TED talks and numerous YouTube tutorials for building self-stimulation devices from household items. The physicist Stephen Hawking even raised the question of whether electrical brain stimulation might lead us to the next step of human evolution (Hawking, 2013).

A rapidly increasing (Minarik et al., 2016) number of studies have investigated to what extent these hopes are realistic and tES is actually effective. The results have been summarized in several meta-analyses, but the results of the single studies as well as of the meta-analyses are heterogeneous. One recent meta-analysis analyzed the effects of single-session tDCS on 30 physiological measures (Horvath et al., 2015a). Only the effects of tDCS on motor-evoked potentials reached significance, whereas none of the other 29 variables was reliably affected by tDCS. A second meta-analysis by the same authors investigated tDCS effects on cognitive outcomes, including executive functions, language production, and memory in healthy adults (Horvath et al., 2015b). None of the 59 comparisons was statistically significant. In contrast, more positive evidence comes from a meta-analysis of tDCS effects on language-related cognitive processes, which found a medium strong effect size (Cohen's d = 0.445; Price et al., 2015). However, with 8 effect sizes from a total of 119 participants, the database was small and the precision of estimation was low as indicated by the large confidence interval (95% CI [0.176, 0.715]). Three meta-analyses focusing on working memory functioning and including between 8 to 36 effect sizes found only weak and inconsistent effects of tDCS (Hill, Fitzgerald, & Hoy, 2016; Mancuso, Ilieva, Hamilton, &

Farah, 2016; Russowsky Brunoni & Vanderhasselt, 2014). The effect sizes were largely unrelated to current density and stimulation duration (Hill et al., 2016) and partly suffered from publication bias (Mancuso et al., 2016). In sum, the meta-analyses so far found mostly low overall effect sizes. However, the heterogeneity of the effect sizes was high, raising the question whether unaccounted variables might moderate the effect of tES on the outcome measures.

In the present meta-analysis, we tested the hypothesis that tES improves learning, that is, the encoding of new information in long-term memory, stronger than processing or recall of already known information from long-term memory. A meta-analytic comparison of these two conditions is possible, because there are basically two types of studies in research on tES. In one type, a psychological construct (e.g., mathematical competence or working memory capacity) is assessed and the participants receive brain stimulation before or during the assessment. In the following, we call this approach *stimulation of test performance*. In the other type of studies, the participants first participate in a learning intervention, for example, they memorize words in a foreign language or practice mental arithmetic. They receive the brain stimulation before or during this learning phase. After the learning phase, the participants complete a learning outcome measure (e.g., assessing how many new words they had learned or how strongly their mental arithmetic competence improved). Usually, there is no separate brain stimulation directly before or during the learning outcome test. In the following, we call this approach *stimulation of learning*.

There are some findings supporting the view that a stimulation of learning might be more effective than the stimulation of test performance. First, anodal tDCS has been shown to reduce regional levels of the inhibitory neurotransmitter *gamma*-Aminobutyric acid (GABA; Stagg et al., 2009). Levels of this neurotransmitter correlate negatively with learning (Floyer-Lea, Wylezinska, Kincses, & Matthews, 2006). Second, anodal tDCS seems to improve learning through neuroplastic alterations of synaptic connections, which share some

similarities with long term potentiation and depression (Paulus, Nitsche, & Antal, 2016; Stagg & Nitsche, 2011). In animal studies, for instance, anodal tDCS promoted long-term potentiation in the motor cortex by increasing the secretion of the brain-derived neurotrophic factor (BDNF), which is an important growth factor in synaptic learning (Castillo, Chiu, & Carroll, 2011; Fritsch et al., 2011). Similar mechanisms might underlie tES effects on human learning in motor tasks and cognitive tasks (B. Krause & Cohen Kadosh, 2013). Finally, the only previous meta-analysis taking this potential moderator into account found a significant effect of tES for the stimulation before or during working memory trainings, that is, learning, but no significant effect of tES for the (online or offline) stimulation before or during working memory assessments (Mancuso et al., 2016). However, the generalizability of the finding is unclear. It is limited to a specific working memory function only, is based on a mere ten effect sizes, and had a fail-safe *N* of seven, meaning that the effect would cease to be significant if only seven unpublished studies with null results would be included in the analyses.

For these reasons, we conducted a meta-analysis designed to provide more comprehensive and generalizable evidence for the effects of tES during learning and test performance. To reduce the heterogeneity of the included studies and to aid the interpretation of the results, we included only studies in the domains of mathematics and language, such as vocabulary learning, sentence comprehension, or arithmetic. Mathematics and language competence are prerequisites of academic achievement in many subjects, professional success, and participation in society (OECD, 2001). Our main research questions were: (1) Does tES affect mathematical and language competence? Based on previous meta-analyses (Hill et al., 2016; Mancuso et al., 2016; Price et al., 2015; Russowsky Brunoni & Vanderhasselt, 2014), we hypothesized to find a positive effect. (2) Does the stimulation before or during a learning phase have stronger effects than the stimulation before or during test performance? We had the hypothesis that this is the case, because tES has been found to

affect processes underlying long-term potentiation, that is, learning (Fritsch et al., 2011; B. Krause & Cohen Kadosh, 2013). (3) Is anodal but not cathodal stimulation effective in the subset of studies using tDCS? We hypothesized this to be the case, as it had been established in a previous meta-analysis on tES (Jacobson, Koslowsky, & Lavidor, 2012). (4) Are the tES effects dosage-specific? We expected to find a systematic relation between stimulation dosage and learning outcomes, which would demonstrate the systematicity of the effects included in the meta-analysis.

5.3 Methods

5.3.1 Inclusion Criteria

Studies were included in the meta-analysis when they fulfilled each of the following criteria: (1) The study used any kind of transcranial electrical brain stimulation (i.e., tDCS, HDtDCS, tACS, tPCS, or tRNS, for more details see section Data Coding). (2) The study included at least one measure of mathematical or language competence, for example, mental arithmetic or vocabulary learning. We included effect sizes from any task assessing domain-specific fact knowledge, skills, procedural knowledge, declarative knowledge, or problem solving as a competence measure. Studies using numerical or verbal stimuli were excluded when they included not directly practically relevant tasks (e.g., the Stroop task), assessed basic cognitive functions (e.g., categorization), or domain-general cognitive processes (e.g., executive functions). (3) The study reported information that allowed us to compute a standardized effect size (e.g., group means, standard deviations, and number of participants). If the necessary information was not reported in a publication, we contacted the authors. If the authors gave the missing information, the study was included. (4) The effect size was collected with a sample size greater than three so that we could calculate the respective variance of the effect size (see Formula 8). (5) The majority of the participants was either healthy or had dyslexia, dyscalculia, or math anxiety but did not suffer from any other

diagnosed neurological or physical disorders. (6) The study was published in the English language.

5.3.2 Literature Search

The literature search is visualized in Figure 1. We searched title, abstract, heading, table of contents, key concepts, original title, tests, and measures of all articles in the electronic database PsycINFO in May 2016. The search included articles from peer-reviewed journals as well as grey literature. The search terms included the types of electrical brain stimulation and mathematical or language competencies. The exact search was conducted in the database PsycINFO with the following search string: ((tDCS or tACS or tPCS or tRNS or transcranial direct current stimulation or transcranial alternating current stimulation or transcranial random noise stimulation or transcranial pulsed current stimulation or brainstimulation) and (math* or num* or arith* or magnitude* or language or linguistic* or comprehension or vocab*)). We also conducted an exploratory internet search and used several mailing lists of relevant research communities to request any unpublished studies or published studies not yet included in the search results. Only one person responded to this request and sent four additional effect sizes with the respective values of the moderator variables.

A trained coder excluded studies with titles and abstracts not fulfilling the inclusion criteria. Subsequently, forty-two full-text articles were obtained for further examination by two trained and independent raters. The inter-rater agreement (100% * *number of agreements* / *number of all full-texts*) for the inclusion of full-texts was 83%. Disagreements were resolved by discussion. All studies except one, authored by Sarkar, Dowker, and Cohen Kadosh (2014), tested healthy participants. This particular study did not find a significant effect regarding brain stimulation on mathematical achievement for participants with math anxiety. To maximize the interpretability of the results, we excluded the study. A total of 35 studies could finally be included in the meta-analysis (see Appendix A and Appendix B for more details).

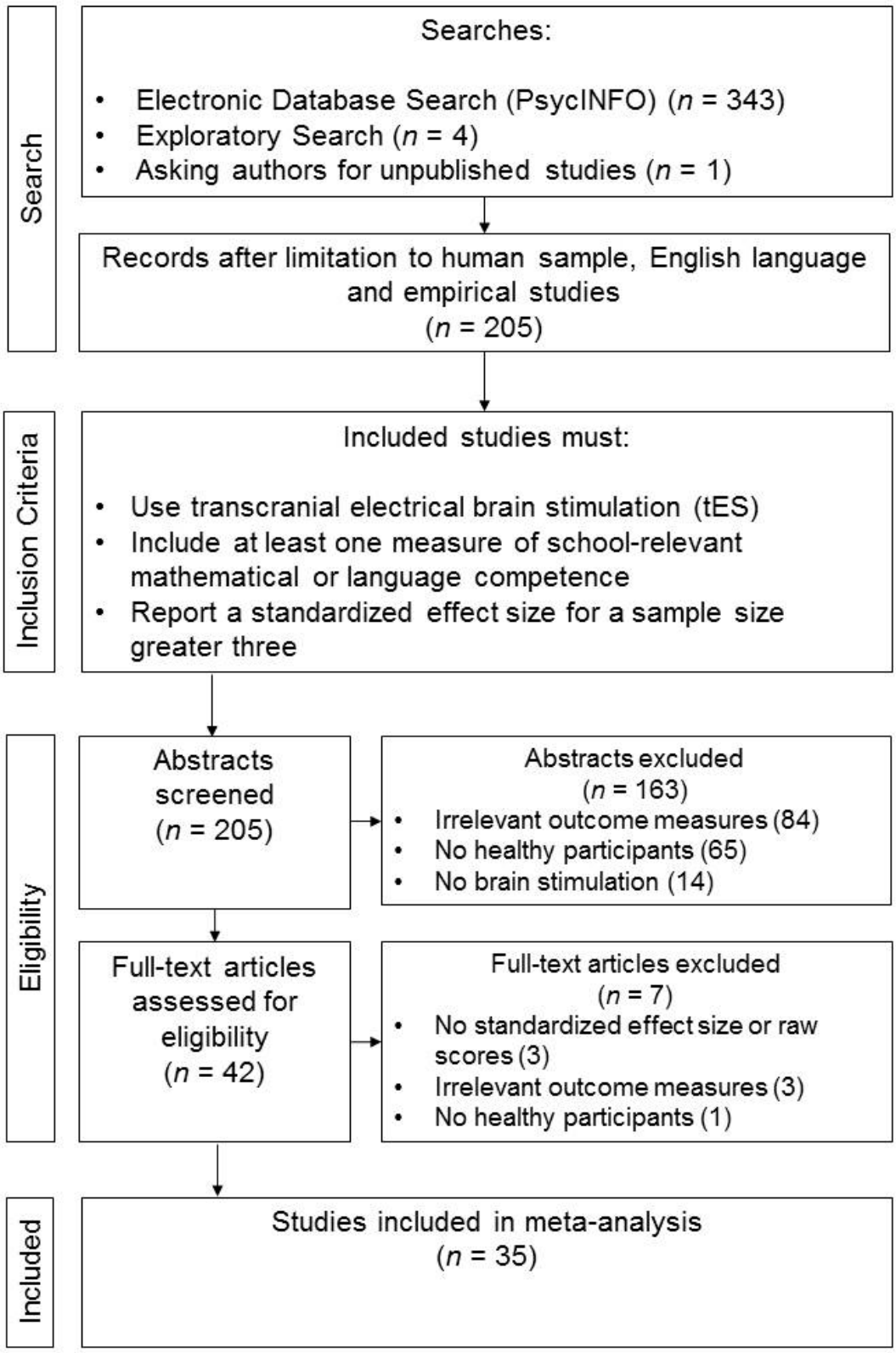

Figure 1. Flow chart of the literature search.

5.3.3 Data Coding

We coded all effects reported in the included studies that resulted from a comparison of a brain-stimulation condition with a no-stimulation control condition (SHAM) or from a pre-test/post-test comparison involving stimulation. We did not limit the number of effects included from each study because studies using several groups, competence measures, or measurement points frequently reported several relevant comparisons. To maximize comparability over the studies, we coded the raw data (i.e., the means and standard deviations of the competence measures) instead of the reported effect sizes and then computed a standardized effect size the same way for each study, as described in the next subsection.

For each included effect, the values of the moderator analyses were coded. An overview of all moderators with their levels and the numbers of studies and effect sizes on each level is given in Table 1. We included moderators relating to the stimulation (stimulation type, timing of stimulation, active electrode, hemisphere of the electrodes, area of the electrodes, size of the electrodes, amperage, current density, number of stimulation sessions, duration of stimulation sessions, time between stimulation sessions), the competence test (time from learning phase, test domain, stimulus novelty, stimulus type, task type, measure), the sample (age, gender, university students), and the study (publication type, research group, randomization, control group, comparison).

In coding the stimulation type, we distinguished among the five most common types of transcranial electrical brain stimulation (tES), namely tDCS, HDtDCS, tACS, tPCS, and tRNS (Nitsche & Paulus, 2000). Including all types of tES in our analysis allowed us to test whether the stimulation method moderates the effect on learning outcomes. In transcranial direct current stimulation (tDCS), a low-intensity (0.5 to 2 mA) constant current is used. The direction of current is determined by the active electrode, which can either be the anode or the cathode. Previous results suggest that anodal stimulation increases cortical excitability (Boros, Poreisz, Munchau, Paulus, & Nitsche, 2008; Nitsche et al., 2003) and cathodal stimulation

decreases cortical excitability (Ardolino, Bossi, Barbieri, & Priori, 2005). High-definition tDCS (HDtDCS) works in a similar way but makes use of smaller gel-based electrodes instead of large sponges. For the constant-current stimulation (tDCS and HDtDCS), we coded the electrode characteristics (active electrode, size and density of the electrode, hemisphere, area) separately for the anode and the cathode. Transcranial alternating current stimulation (tACS) is also similar to tDCS but uses a low-intensity (0.25 – 1 mA) bidirectional, biphasic current and can be applied at different frequencies. Transcranial pulsed current stimulation (tPCS) uses repeated bursts of low-intensity (0.6 to 1 mA) pulses. Transcranial random noise stimulation (tRNS) employs very low-intensity (-500 to 500 μA) alternating current with random amplitudes and frequencies (Moreno-Duarte et al., 2014). Implementation of tES may further vary in parameters of dosage, describing the amount of current delivered (in mA), the size of the electrodes (in cm^2), the resulting average density of the current (in mA/cm^2), the duration of the stimulation (in minutes), and the placement of the electrode. It is suggested that difference in the stimulation dosage may lead to changes and inversion of the direction of effects of tES (Moreno-Duarte et al., 2014).

All competence test means, *SD*s, and moderator variables were independently coded by two trained raters. The inter-rater agreement (100% * number of agreeing values / number of all coded values) for the coded variables ranged from 72% to 100%, with an average of 93%. Disagreements were resolved by discussion. If relevant information was missing or was reported in an ambiguous way, we contacted the authors via e-mail.

Table 1

Summary of coded characteristics, number of coded studies (*j*), and effect sizes (*k*)

	No. of studies *j*	No. of effect sizes *k*
Overall	35	246
Stimulation		
Timing of stimulation		
Test performance	25	186
Learning	9	51
Both	4	9
Stimulation Method		
tDCS	29	213
HDtDCS	1	8
tACS	0	0
tPCS	1	2
tRNS	4	23
Active Electrode		
Anodal stimulation	24	167
Cathodal stimulation	11	41
Alternating current	4	23
Hemisphere - anode		
Left	28	168
Right	14	44
Right and left	4	20
Central	1	4
Hemisphere - cathode		
Left	16	74
Right	22	108
Right and left	4	20
Central	3	18
Area of stimulation - anode		
Frontal cortex	14	74
Central cortex	3	10
Parietal cortex	16	98
Temporal cortex	2	20

Occipital cortex	2	14
Supraorbital cortex	6	17
Shoulder	2	7
Area of stimulation - cathode		
Frontal cortex	12	53
Central cortex	5	24
Parietal cortex	9	33
Temporal cortex	1	2
Occipital cortex	1	12
Supraorbital cortex	10	72
Ear	1	1
Cheek	1	12
Shoulder	3	23
Number of stimulation sessions	16	98
Duration of stimulation/session [min]	15	96
Current density anode [mA/cm^2]		
$0 \leq J < .04$	12	86
$.04 \leq J < .08$	19	93
$J \geq .08$	9	65
Current density cathode [mA/cm2]		
$0 \leq J < .04$	14	91
$.04 \leq J < .08$	18	92
$J \geq .08$	7	41
Electrode size - anode [cm^2]	33	243
Electrode size - cathode [cm2]	31	223
Electrode sizes		
Same	23	144
Different	9	79
Competence test		
Domain		
Mathematics	17	85
Language	18	161
Measure		
Solution accuracy	25	121
Solution time	21	109

Other	5	16
Stimulus novelty		
Familiar	29	205
Unfamiliar	3	30
Stimulus type		
Pictures	8	71
Numbers	12	58
Words	5	55
Words and pictures	4	46
Other	3	5
Task type		
Picture naming	5	54
Sentence comprehension	2	27
Vocabulary	2	26
Fact recall	2	13
Verbal fluency	3	11
Grammar	1	2
Other language task	4	29
Mental arithmetic	7	33
Vocabulary	2	26
Magnitude comparison	5	24
Number line estimation (number to space)	2	13
Other mathematics task	5	14
Time from learning phase	6	45
Learner Characteristics		
Age [years]	32	231
University students		
Yes	11	63
No	7	31
Gender [%]	32	225
Study characteristics		
Randomization		
Yes	26	163
No	4	37
Control group		
Sham	34	232

Nothing	1	2
No control	1	11
Comparison		
Pre-post gains	12	64
Post-post	24	171
Pre-post	1	11
Publication type		
Peer-reviewed	34	242
Unpublished	1	4
Research group		
Cohen Kadosh	5	15
Grabner	4	27
Other	26	204

5.3.4 Data Preparation

For each effect, we computed the effect size as Cohen's *d* from the coded raw data using syntax. For one study (Flöel, Rösser, Michka, Knecht, & Breitenstein, 2008) the raw data was not available. Instead, we included two *t*-values for dependent measures reported in that study and transformed them into Cohen's *d*. In the following, we present the formulas used for the calculation of Cohen's *d* that were derived from various sources (Borenstein, 2009; Jacob Cohen, 1988; Morris, 2008; Morris & DeShon, 2002). We use the term *treatment* to refer to the stimulation (i.e., tES) and we use the notation *M* for Mean, *SD* for standard deviation, *T1* for the pre-test, *T2* for the post-test, and *n* for sample size.

For studies reporting pre-post comparisons for a treatment condition, we computed the effect sizes as the following:

$$d = \frac{M_{Treatmen,T2} - M_{Treatment,T1}}{SD_{Treatment,T1}} \quad (1)$$

To calculate effect sizes from a post-post comparison between a treatment group and a control group at T2, we used the following formula:

$$d = \frac{M_{Treatment,T2} - M_{Control,T2}}{\sqrt{\frac{(n_{Treatment,T2}-1)\times SD^2_{Treatment,T2} + (n_{Control,T2}-1)\times SD^2_{Control,T2}}{n_{Treatment,T2}+n_{Control,T2}-2}}} \quad (2)$$

If the study reported descriptive data for a treatment condition and a control condition at T1 and T2 which allows calculation of pre-post gains, we calculated effect sizes for within-subjects designs as the following:

$$d = \frac{(M_{Treatment,T2} - M_{Treatment,T1}) - (M_{Control,T2} - M_{Control,T1})}{\frac{SD_{Treatment,T1} + SD_{Control,T1}}{2}} \quad (3)$$

For effects on pre-post gains for between-subjects designs, we used the following formula:

$$d = \frac{(M_{Treatment,T2} - M_{Treatment,T1}) - (M_{Control,T2} - M_{Control,T1})}{\sqrt{\frac{(n_{Treatment,T1}-1)\times SD^2_{Treatment,T1} + (n_{Control,T1}-1)\times SD^2_{Control,T1}}{n_{Treatment,T1}+n_{Control,T1}-2}}} \quad (4)$$

When only *t*-values were reported, we transformed the *t*-values (*t*) to *d* for within designs as follows:

$$d = \frac{t}{\sqrt{n}} \quad (5)$$

We transformed t-values for between designs with:

$$d = t \times \sqrt{\frac{n_{Treatment} + n_{Control}}{n_{Treatment} \times n_{Control}}} \quad (6)$$

When studies reported *F*-values for univariate ANOVAs we transformed the *F*-values to Cohen's *d* by:

$$d = \sqrt{F \times \frac{n_{Treatment} + n_{Control}}{n_{Treatment} \times n_{Control}}} \quad (7)$$

After computing Cohen's *d*, we used the following equation to compute the variance of each effect size (F. L. Schmidt & Hunter, 2015):

$$Var(d) = \left(\frac{N-1}{N-3}\right)\left(\frac{4}{N}\right)\left(\frac{1+d^2}{8}\right) \quad (8)$$

A problem in averaging the effect sizes found with tDCS is that anodal stimulation tends to enhance cognition whereas cathodal stimulation might inhibit cognition or have no effect (Jacobson et al., 2012). Averaging over these positive and negative effects would incorrectly result in an effect size close to zero. To avoid this problem, we recoded the signs of all effect sizes so that lower competence in a cathodal-simulation condition and higher competence in an anodal-stimulation condition had positive signs because they were in line with the expectations. Higher competence in a cathodal-simulation condition and lower competence in an anodal-stimulation condition had negative signs to indicate that they were not in line with the expectations. Thus, a positive overall effect size for tDCS in our meta-analysis indicates that cathodal stimulation decreases competence and/or anodal stimulation increases competence. For other stimulation types, this issue did not arise because there was no constant current.

In meta-analyses, it is generally possible to correct the results for measurement error (F. L. Schmidt & Hunter, 2015). However, because only one of the included studies reported the reliability of the outcome measure, we could not correct for measurement errors in our meta-analysis. We detected and removed outlier effect sizes (effect sizes with a *z*-score greater than 3.29; Tabachnick & Fidell, 2014) before the meta-analytic aggregation. There were only two outliers.

5.3.5 Statistical Analysis

Most of the included studies had more than one relevant effect size. Effect sizes from the same study are statistically dependent and thus cannot be handled by conventional meta-analytic techniques (Hedges et al., 2010). Therefore, we employed robust variance estimation (RVE; Hedges et al., 2010; Tanner-Smith & Tipton, 2014; Tanner-Smith et al., 2016), because it allows for the inclusion of statistically dependent effect sizes in a meta-analysis

without requiring information about the inter-correlation between effect sizes within studies. Given the expected heterogeneity, we used random-effects models for all analyses (Raudenbush, 2009). Mean effect sizes and meta-regression models using robust variance estimation were estimated using a weighted least squares approach (Hedges et al., 2010; Tanner-Smith & Tipton, 2014). To estimate the overall effect of tES on mathematical and language competence, we estimated a simple RVE meta-regression model, as follows:

$$y_{ij} = \beta_0 + u_j + e_{ij} \quad (9)$$

where y_{ij} is the *i*th effect size in the *j*th study, β_0 is the average population effect, u_j is the study level random effect such that $\text{Var}(u_j) = \tau^2$ is the between-study variance component, and e_{ij} is the residual for the *i*th effect size in the *j*th study. To estimate the variability in the effect size due to moderator variables, we estimated the following mixed-effects RVE meta-regression model:

$$y_{ij} = \beta_0 + \beta_1(Moderator_1)_{ij} + \cdots + \beta_k(Moderator_k)_{ij} + u_j \quad (10)$$

In this model, the moderator represents a continuous variable (e.g., age) or specific dummy-coded levels (e.g., the three levels of current density) of an included moderator variable.

We used the *robumeta* package (Z. Fisher & Tipton, 2014) in the R statistical environment (R Core Team, 2014) for the statistical analyses. To estimate the proportion of explained variance R^2 for the multiple regression models with categorical moderators with more than two levels, we used the following formula (Tabachnick & Fidell, 2014):

$$R^2 = \sum_{i=1}^{k} r_{yi}\beta_i \quad (11)$$

with the bivariate correlation *r* and the regression coefficient β.

A common problem of meta-analyses is publication bias (Hunter & Schmidt, 2004; Rothstein, Sutton, & Borenstein, 2005). We approached this problem by analyzing the symmetry of the distribution around the mean through visual inspection of the funnel plots, computing Egger regressions (Egger et al., 1997), using the trim-and-fill method (Duval &

Tweedie, 2000), and estimating HC intervals, representing confidence intervals with robustness to publication bias (Henmi & Copas, 2010).

5.4 Results

5.4.1 Characteristics of the Included Studies

A total of 35 studies reporting 246 effect sizes obtained with 885 participants were included in the meta-analysis. The studies were published between 2008 and 2016. Of the reported effect sizes, 65% related to language and 35% related to mathematics. The studies used various language tasks (22% picture naming, 11% sentence comprehension, 11% vocabulary 5% fact recall, 5% verbal fluency, 12% other behavioral competence measures) or mathematical tasks (14% mental arithmetic, 10% magnitude comparison, 5% number line estimation, 6% other behavioral competence measures). They were also representative of the research literature in that they included wide ranges of stimulation types, locations, timings, current densities, and numbers of stimulation sessions.

5.4.2 Overall Effectiveness of tES

The meta-analytic results are listed in Table 2. The overall result supported the hypothesis that tES positively affects mathematical and language competence. The overall Cohen's *d* was 0.343. The 95% confidence interval [0.173, 0.513] did not include the zero, thus indicating that the effect size was significantly different from zero with $p < .05$. The I^2 of 51.8 indicated a moderate heterogeneity of the effect sizes (J. P. T. Higgins & Thompson, 2002) implying that the effectiveness of tES is moderated by third variables.

Table 2

Number of studies (*j*), number of effect sizes (*k*), effect size (*d*), 95% confidence interval, measure of heterogeneity τ^2 for the levels of the moderator variables, and significance and R^2 for the moderator analyses

	j	*k*	*d*	*95% CI*	τ^2	Moderator	
						Sign.	R^2
Overall	35	246	0.343	[0.173, 0.513]	.200		
Stimulation							
Time of Stimulation						**	.163
Test performance	25	186	0.207	[0.021, 0.393]	.197		
Learning	9	51	0.712	[0.387, 1.040]	.035		
Both	4	9	0.763	-[a]	-[a]		
Stimulation method						-	-
tDCS	29	213	0.267	[0.100, 0.434]	.171		
tRNS	4	23	1.190	-[a]	-[a]		
Active electrode						*ns*	.033
Anodal stimulation	24	167	0.343	[0.143, 0.543]	.005		
Cathodal stimulation	11	41	0.053	[-0.150, 0.256]	.005		
Alternating current	4	23	1.190	-[a]	-[a]		
Hemisphere - anode						*ns*	.022
Left	28	168	0.395	[0.191, 0.589]	.215		
Right	14	44	0.177	[0.016, 0.337]	.116		
Right and left	4	20	0.406	-[a]	-[a]		
Hemisphere - cathode						*	.063
Left	16	74	0.111	[-0.043, 0.264]	.118		
Right	22	108	0.407	[0.163-0.650]	.252		
Right and left	4	20	0.406	-[a]	-[a]		
Central	3	18	0.517	-[a]	-[a]		
Area of stimulation - anode						*ns*	.017
Frontal cortex	14	74	0.353	[0.020, 0.687]	.229		
Central cortex	3	10	0.320	-[a]	-[a]		
Parietal cortex	16	98	0.385	[0.117, 0.654]	.120		
Temporal cortex	2	20	0.184	-[a]	-[a]		

Occipital cortex	2	14	-0.140	-[a]	-[a]		
Supraorbital cortex	6	17	0.186	[-0.130, 0.504]	.010		
Right Shoulder	2	7	-0.188	-[a]	-[a]		
Area of stimulation - cathode						*ns*	.088
Frontal cortex	12	53	0.169	[-0.125, 0.463]	.116		
Central cortex	5	24	0.468	-[a]	-[a]		
Parietal cortex	9	33	0.325	[-0.092, 0.742]	.148		
Supraorbital cortex	10	72	0.451	[0.006, 0.897]	.371		
Right Shoulder	3	23	0.103	-[a]	-[a]		
Number of stimulation sessions	16	98	-	-	-	**	.095
Duration of stimulation/session [min]	15	96	-	-	-	*ns*	.146
Current density anode [mA/cm^2]						**	.099
$0 \leq J < .04$	12	86	0.329	[0.150, 0.635]	.026		
$.04 \leq J < .08$	19	93	0.434	[0.141, 0.728]	.368		
$J \geq .08$	9	65	0.041	[-0.143, 0.225]	.000		
Current density cathode [mA/cm^2]						**	.094
$0 \leq J < .04$	14	91	0.360	[0.124, 0.596]	.096		
$.04 \leq J < .08$	18	92	0.455	[0.158, 0.751]	.282		
$J \geq .08$	7	41	0.068	[-0.201, 0.337]	.000		
Amperage [mA]						*	.054
1.0	16	103	0.481	[0.212, 0.750]	.102		
1.5	7	34	0.445	[-0.249, 1.140]	.598		
2.0	12	108	0.145	[-0.006, 0.356]	.133		
Electrode size - anode [cm^2]	33	243				*ns*	.001
Electrode size - cathode [cm2]	31	223				*ns*	.012
Electrode sizes						*ns*	.001
Same	23	144	0.457	[0.201, 0.713]	.300		
Different	9	79	0.232	[-0.029, 0.493]	.065		
Competence test							
Domain						*ns*	.004
Mathematics	17	85	0.289	[0.040, 0.538]	.120		
Language	18	161	0.395	[0.139, 0.652]	.311		

Measure						*ns*	.020
Solution accuracy	25	121	0.280	[0.043, 0.516]	.219		
Solution time	21	109	0.260	[0.036, 0.485]	.210		
Other	5	16	0.488	-[a]	-[a]		
Stimulus novelty						-[a]	-[a]
Familiar	29	205	0.301	[0.106, 0.495]	.199		
Unfamiliar	3	30	0.593	-[a]	-[a]		
Stimulus type						*ns*	.020
Pictures	8	71	0.457	[0.131, 0.783]	.393		
Numbers	12	58	0.218	[-0.092, 0.528]	.121		
Words	5	55	0.066	-[a]	-[a]		
Words and pictures	4	46	0.362	-[a]	-[a]		
Other	3	5	0.925	-[a]	-[a]		
Task type						-[a]	-[a]
Picture naming	5	54	0.171	-[a]	-[a]		
Sentence comprehension	2	27	0.009	-[a]	-[a]		
Vocabulary	2	26	0.652	-[a]	-[a]		
Fact recall	2	13	0.892	-[a]	-[a]		
Verbal fluency	3	11	0.525	-[a]	-[a]		
Other language task	4	29	0.657	-[a]	-[a]		
Vocabulary	2	26	0.652	-[a]	-[a]		
Mental arithmetic	7	33	-0.014	[-0.146, 0.118]	.000		
Magnitude comparison	5	24	0.352	-[a]	-[a]		
Number line estimation	2	13	0.470	-[a]	-[a]		
Other mathematics task	5	14	0.680	-[a]	-[a]		
Time from learning phase	6	45				*ns*	.000
Learner Characteristics							
Age [years]	32	231				*ns*	.001
University students						*ns*	.016
Yes	11	63	0.229	[-0.016, 0.474]	.176		
No	7	31	0.410	[-0.099, 0.920]	.161		
Gender [%]	32	225				*ns*	.006
Study characteristics							
Randomization						-[a]	-[a]
Yes	26	163	0.391	[0.180, 0.603]	.241		
No	4	37	0.284	-[a]	-[a]		

Comparison						*ns*	.062
Pre-post gains	12	64	0.366	[0.059, 0.673]	.267		
Post-post	24	171	0.365	[0.144, 0.587]	.183		
Research group						-[a]	-[a]
Cohen Kadosh	5	15	0.727	-[a]	-[a]		
Grabner	4	27	0.217	-[a]	-[a]		
Other	26	204	0.299	[0.111, 0.487]	.203		

[a] insufficient number of data points for the analysis

Note. ns: not significant; * $p < .05$, ** $p < .01$; moderators without levels were used as continuous predictors in meta-regressions; all moderators with only one study were omitted from this table.

5.4.3 Moderator Analyses

The results also supported our main hypothesis, that tES effects are stronger for the stimulation of learning than for the stimulation of test performance. The effect size was d = 0.211 for the stimulation of test performance. With d = 0.712, the effect was three times stronger for the stimulation in a learning phase. A stimulation in both phases was descriptively even stronger with d = 0.763. All three effect sizes are statistically significant, but it should be noted that the lower bound of the confidence interval of the effect size for the stimulation of test performance is close to zero. Whether learning or test performance was stimulated had a highly significant effect ($p < .01$), explained a variance proportion of .163 of the effect sizes, and was the strongest moderator in the meta-analysis.

Further moderator analyses showed that the effect of tES significantly varied under different stimulation conditions. Descriptively, the effect was much stronger for tRNS than for tDCS, but the small number of studies with tRNS combined with the non-independence of many effect sizes did not permit testing the statistical significance of this difference. In tDCS and HDtDCS, a significant effect was only found for anodal but not for cathodal stimulation. Anodal stimulation was effective with d = 0.343, whereas cathodal stimulation did not

influence learning outcomes with d = 0.053. Similarly, hemispheres and areas of the stimulation differed descriptively in their mean effect sizes, but these differences reached significance only for the hemisphere of the cathode. When the cathode was on the right hemisphere, effect sizes were significantly larger as compared to when it was placed on the left hemisphere.

The effect of tES was dosage-specific. The effect sizes increased with the number of stimulation sessions and were highest for a current density between 0.04 and 0.08 mA/cm^2 and a current of 1 mA. The effect of tES was significant for mathematical as well as for language measures and for solution accuracy as well as solution time. In line with our hypothesis that the stimulation of learning is more effective than the simulation of test performance, the effect size was almost twice as high for studies presenting the participants with novel stimuli instead of familiar ones. The descriptive differences between task types should not be interpreted, because there were only few studies using each type of task so that task type might be confounded with other study characteristics in our analyses. For the learner and study characteristics, no significantly moderator effects were found.

5.4.4 Publication Bias

There was no publication bias on the study level and only a weak publication bias on the effect size level (see Figure 2). Egger regressions for random-effects (Egger et al., 1997) did not indicate asymmetry on the study level (Z = 1.585, p = .113) but did indicate asymmetry on the effect size level (Z = 4.083, p < .0001). The trim-and-fill method did not add any cases to the distribution to make it more symmetric. This indicates that the publication bias was too small to be detected with the trim-and-fill method. In line with this, the estimation of the HC intervals (Henmi & Copas, 2010) indicated only a weak publication bias and reduced the effect size slightly from d = 0.328, 95% CI [0.169, 0.488] to d = 0.314, 95% CI [0.150, 0.478] on the study level and from d = 0.276, 95% CI [0.194, 0.358] to d = 0.238, 95% CI [0.166,

0.310] on the effect size level. This shows that the positive findings cannot entirely be attributed to a publication bias.

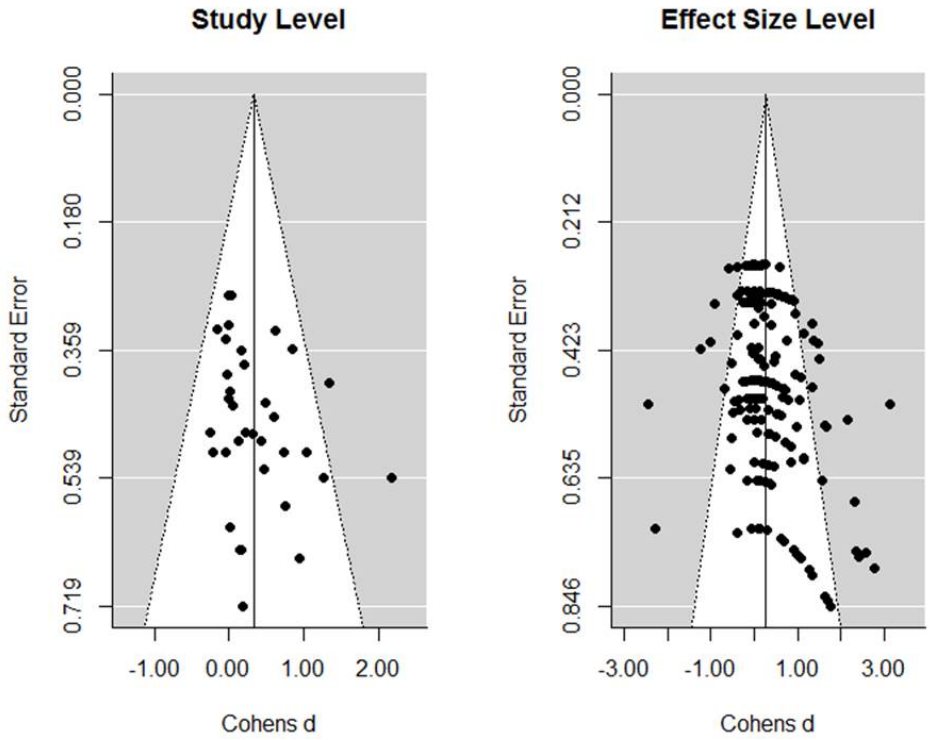

Figure 2. Funnel plots for average effect size per study (left) and all effect sizes (right).

5.5 Discussion

5.5.1 Effects of tES on Practically Relevant Learning Tasks

The present meta-analysis aimed to investigate the effects of brain stimulation on practically relevant learning tasks and whether the effectiveness differs due to timing of the stimulation (i.e., whether it was provided in the learning or testing phase) and characteristics of the stimulation (i.e., anodal vs. cathodal). In line with our expectations, tES had a positive overall effect on learning outcomes. Further, tES in a learning phase, where participants had to encode new information, had a much stronger effect than tES of test performance, where participants had to recall information. This may be explained by neuroplastic alterations of synaptic connections via different pathways (e.g., BDNF secretion, GABAergic activity; Castillo et al., 2011; Fritsch et al., 2011; Paulus et al., 2016; Stagg & Nitsche, 2011).

Elaborating the details of these mechanisms in cognitive learning remains an important topic for future research.

Brain stimulation was effective for anodal but not for cathodal stimulation. The findings are in line with previous meta-analytic results, indicating effectiveness of anodal but not cathodal stimulation (Jacobson et al., 2012). Research investigating the mechanisms of brain stimulation found evidence that anodal stimulation increases cortical excitability (Boros et al., 2008; Nitsche et al., 2003) and cathodal stimulation decreases cortical excitability (Ardolino et al., 2005) through modulation of membrane potentials. Following, firing threshold of neurons decrease during anodal stimulation and increase during cathodal stimulation, so that neurons in the stimulated area require less input to fire or become inhibited and require more input, respectively. These mechanisms are also relevant for learning, which likely explains the present results. By averaging over (effective) anodal and (ineffective) cathodal stimulation in our meta-analysis, we obtained lower effect sizes than we would have obtained by focusing on anodal stimulation only. We still think that including anodal and cathodal stimulation in the current analyses is warranted, because the differential effects of anodal and cathodal stimulation are not well understood yet.

The effect of tES was dosage-specific. The effectiveness of tES increased with the number of stimulation sessions and was associated with the amperage and current density of the stimulation. tES was most effective for a current of 1 mA and a current density between 0.04 and 0.08 mA/cm^2. The findings therefore demonstrate that characteristics of the application of tES are essential for the effectiveness of the method.

The finding that anodal tES enhances learning mathematics and language is in line with studies suggesting that anodal tES might also have positive effects on motor learning (Ammann, Spampinato, & Márquez-Ruiz, 2016; Buch et al., 2017). However, we do not know of any direct comparisons of the stimulation of academic learning versus motor learning. Future stimulation studies will have to systematically investigate the commonalities

and differences of academic and motor learning in terms of effective stimulation methods, stimulated brain regions, and underlying neurophysiological processes.

5.5.2 Generalizability of the Findings

In sum, the results of the present study support the view that tES affects performance and learning across a variety of competence measures and activated brain regions. Since most included studies were randomized trials involving a SHAM control condition, the findings can be interpreted in terms of a causal effect. Several findings demonstrate the validity and systematicity of the meta-analytic results. Direct checks indicated only a very weak publication bias, the effects of tES were dosage-specific, and were consistently found for mathematics and language learning, with accuracy and speed measures. We were also able to replicate the previous meta-analytic finding that, in direct current stimulation, anodal but not cathodal stimulation has positive causal effects. Despite the high statistical power that comes with 246 effect sizes, there was no evidence that our results would be specific to only mathematics learning or language learning, to only accuracy measures or speed measures, to only women or men, or to only university students or other participants. This demonstrates the consistency and generalizability of the findings.

Our literature search and inclusion criteria were designed to also include adequate studies on tES effects on mathematical and language learning in samples with dyslexia, dyscalculia, or math anxiety. As there was only one study matching the inclusion criteria and using a clinical sample, we excluded the study from the meta-analysis. The generalizability of our results is thus limited to healthy participants only.

5.5.3 Implications for Research and Practice

Overall, the results of the present study imply that previous meta-analyses might have dramatically underestimated beneficial effects of tES on cognition and performance, because they did not differentiate between the stimulation of learning and the stimulation of test performance. Most empirical studies so far investigated the latter, but as we show here, the

effect sizes are more than three times higher in the former case. This is a highly relevant finding for the research field of educational neuroscience, in which tES has been intensively discussed as potential future means to support knowledge acquisition in individuals with learning difficulties, especially in language and mathematics (Iuculano & Cohen Kadosh, 2014; B. Krause & Cohen Kadosh, 2013). The results give evidence that despite the heterogeneity in competence measures and activated brain regions which is accompanied by significant noise, the signal was strong enough to show significant results through this noise.

These highly encouraging findings do not imply that tES should be used in practical educational contexts right away. This would be a grave mistake, both from a scientific and practical perspective. From a scientific point of view, there is reason for concern because of the low reproducibility of empirical studies in psychology and related fields (Open Science Collaboration, 2015). These are partly because of questionable research practices, for example, using several learning outcome measures in an intervention study but reporting only the ones that showed a significant effect. So far, this problem has not been explicitly discussed in research on tES. If present in several studies, the problem could have biased the results of our meta-analysis. Therefore, the primary implication of the current findings is not that tES demonstrated effectivity, but rather that the first phase of tES research with its small sample sizes and explorative approaches should come to an end and be followed by a second phase characterized by replication attempts, preregistered trials, larger sample sizes, open data, a full documentation of rigorous research practices, and further meta-analytic integration.

From a practical perspective, there is reason for caution because the present results had exclusively been obtained in laboratory studies with healthy adults. The generalizability of the findings to learning in classroom settings and to children and adolescents, especially those suffering from learning difficulties, is unclear, as only few studies exist (e.g., Costanzo et al., 2016; Looi et al., 2017). In addition, almost nothing is currently known about the cognitive

side effects of stimulation, that is, whether improving one brain function can unintentionally impair other brain functions (e.g., Iuculano & Cohen Kadosh, 2013; Sarkar, Dowker, & Cohen Kadosh, 2014). When tES is used, safety issues have to be carefully considered (Davis, 2014). Self-built stimulation devices are not recommended due to safety concerns and the potential for unintentional impairment of brain functions (Wurzman, Hamilton, Pascual-Leone, & Fox, 2016). The cost–benefit relationship needs to be considered because there is an abundance of instructional methods with lower financial costs and stronger effects on learning than tES (e.g., Hattie, 2009). Finally, there is the need for an ethical debate about cognitive self-enhancement and questions such as accepting or challenging the limits of human nature, freedom of choice for learners, informed consent in children, and the fair distribution of technological learning resources worldwide (Cohen Kadosh, Levy, O'Shea, Shea, & Savulescu, 2012).

5.6 Conclusion

The meta-analytic results demonstrate that studies in which brain stimulation was administered in a learning phase found, on average, higher effect sizes than studies in which brain stimulation was administered in a test performance phase. This is in line with studies showing effects of tES on neurotransmitters and synaptic mechanisms associated with learning, that is, the encoding of new information, rather than recall or recognition. The meta-analytic results were also highly consistent in that the overall effect was stimulation-dosage specific and significant only for anodal stimulation, as had been found in a previous meta-analysis. Yet, despite the highly encouraging results, it is too early to use tES in practical educational settings, because there is still a lack of pre-registered independent replications, field experiments in schools, and studies with children.

6. Study 2: Domain-Specific Prior Knowledge and Learning: A Meta-Analysis

6.1 Abstract

Domain-specific prior knowledge has been hypothesized to be one of the strongest determinants of learning. Most researchers assumed positive effects. Others emphasized that prior knowledge can hinder learning (e.g., misconceptions, the Einstellung effect, or negative transfer). The present meta-analysis integrated the extensive empirical findings on the relation between the amount of domain-specific prior knowledge and subsequent cognitive learning outcomes (i.e., knowledge and achievement). We included 240 articles reporting 4327 effect sizes obtained with 62,129 participants. Almost all studies investigated the relation between prior knowledge and posttest knowledge. Individual differences in prior knowledge were a strong predictor of individual differences in knowledge and achievement after learning (r^+ = .521, 95% CI [.491, .550]). Controlling for intelligence did not reduce this effect size. The effect is partly causal, as it was weaker but still significantly positive for randomized controlled trials. The strongest moderators were the cognitive demands of the instruction, correctness of knowledge, and educational level. The findings demonstrate that prior knowledge relates correlationally, longitudinally and causally to knowledge and achievement after learning. Only very few studies investigated the relation between prior knowledge and knowledge gains from pre-test to posttest. This finding highlights the need for more randomized controlled studies investigating the causal effect of prior knowledge on pretest-posttest learning gains. Prior knowledge is of a high predictive power for cognitive learning outcomes. Considering it can improve the effectivity of learning environments such as school instruction. We present a framework to guide future research on the processes mediating and moderating these processes.

6.2 Theoretical Background

"The most important single factor influencing learning is what the learner knows already". Since Ausubel (1968, p. vi) formulated this hypothesis, many researchers have investigated the effects of prior knowledge. For example, memory researchers demonstrated that the content of long-term memory affects how new information is processed in working memory and encoded in long-term memory (Baddeley et al., 2009). Cognitive scientists devised models of these mechanisms (cf. Anderson et al., 2004; Gopnik & Wellman, 2012; Laird, 2012). Developmental psychologists of the past (Piaget, 1971; Vygotsky, 1978) and present (e.g., Case, 1992; Siegler, 1996; Wellman & Gelman, 1992) investigated the role of knowledge in cognitive development over the lifespan. Educational psychologists incorporated prior knowledge as a central component in theories on academic achievement (Thompson & Zamboanga, 2003), expert performance (Ericsson & Charness, 1994), transfer (Barnett & Ceci, 2002; Singley & Anderson, 1989; Thorndike & Woodworth, 1901), multimedia learning (Mayer, 2001), and conceptual change (diSessa, 2008; Vosniadou, 1994). Surprisingly, this plethora of research has left central questions open: how strong is the overall effect of prior knowledge on learning? How variable are the effect sizes? Are there systematic variations in the effect sizes that can be explained by moderator variables? – To our knowledge, no comprehensive meta-analysis on these questions has been published yet.

For the purpose of the present investigation we define *knowledge* as information stored in memory (e.g., Anderson, 1983; Weinert, 1999). In line with how the term knowledge is typically used in the psychological and educational research literature, this definition includes declarative knowledge about abstract and relational concepts (Goldwater & Schalk, 2016) and more isolated facts (Michael Schneider & Grabner, 2012), as well as procedural knowledge about how to solve problems (Anderson et al., 2004). It also includes scientifically incorrect misconceptions as well as scientifically correct concepts (Shtulman & Valcarcel, 2012; Smith III. et al., 1993). Knowledge is *domain-specific* when its relates to the key principles in a

domain, for example, the concept of equivalence in mathematics or the concept of force in physics (Carey & Spelke, 1994; Wellman & Gelman, 1992). Domain specific knowledge is sometimes also termed content knowledge (Chi & Ceci, 1987) and has been described as a component of academic achievement (OECD, 2016; Steinmayr et al., 2014). We define *prior knowledge* as the knowledge available in a person's long-term memory at the onset of learning (cf. Dochy & Alexander, 1995).

These definitions have two implications for the current study. First, they imply that investigations of the influence of prior knowledge on subsequent learning require the assessment or manipulation of knowledge before learning (a pretest) and the assessment of learning outcomes after learning (a posttest). Only then, differences in prior knowledge before learning allow a prediction of the learning outcomes after learning or the learning gains. Second, the definitions imply that prior knowledge is usually partly or fully incorrect or at least incomplete compared to what is to be learned in a learning phase. Only if this is the case, learning something relevant is still possible, so that the relation between prior knowledge and learning can be investigated. This partly incorrect or incomplete nature of prior knowledge raises the question whether prior knowledge always facilitates or sometimes hinders further learning.

6.2.1 Does Prior Knowledge Facilitate or Hinder Learning?

In the literature on knowledge acquisition, there are two seemingly conflicting hypotheses about the relationship between prior knowledge and subsequent learning – the knowledge-is-power hypothesis and the knowledge-as-burden hypothesis (cf. Dochy et al., 1999; W. Schneider, 1993; Stern, 2001). There exists empirical evidence supporting both hypotheses.

According to the so-called *knowledge-is-power hypothesis* (Hambrick & Engle, 2002), prior knowledge has a positive effect on learning (W. Schneider et al., 1989). Research has suggested several mechanisms that contribute to this effect. For example, prior knowledge

guides the learners' attention (e.g., Tanaka, Kiyokawa, Yamada, Dienes, & Shigemasu, 2008; Yu, Zhong, & Fricker, 2012), facilitates the interpretation and encoding of new information (Kintsch, 1994; van Kesteren et al., 2014), and allows for the bundling of new information into chunks that can efficiently be processed, memorized, and retrieved (Chase & Simon, 1973; Ericsson, Chase, & Faloon, 1980; Gobet et al., 2001). Prior knowledge about the effectivity and efficiency of problem-solving strategies also allows for more exploration, goal-directed behavior, and the construction of more advanced strategies (Michael Schneider, Rittle-Johnson, & Star, 2011; Siegler, 1996).

In contrast, the *knowledge-as-burden hypothesis* predicts negative effects of prior knowledge on learning and transfer (Bilalić, McLeod, & Gobet, 2008; M. Fisher & Keil, 2016; Luchins & Luchins, 1987; Vosniadou, Vamvakoussi, & Skopeliti, 2008). For example, misconceptions as well as incomplete correct knowledge in a domain (e.g., that the surface of the earth looks flat in everyday life) can give rise to incorrect conclusions (e.g., the earth is a disc; Vosniadou & Brewer, 1992). Learners with high correct prior knowledge in a domain tend to pay selective attention to the features of a situation that have been relevant for solving problems in the past. This can induce perceptual biases (Hecht & Proffitt, 1995; Lewandowsky & Kirsner, 2000) or the prevention that learners look for new and better problem solutions (Einstellung effect; Bilalić, McLeod, & Gobet, 2010; Luchins & Luchins, 1959). Having more knowledge elements about a topic increases the probability of intrusions or interferences involving these elements (Arkes & Freedman, 1994; Castel, McCabe, Roedinger, & Heitman, 2007). High procedural prior knowledge is usually acquired through extensive practice, which can lead to efficient automatized but also inflexible routines that hamper adaptive problem solving (A. Johnson, 2003; Müller, 1999). Higher levels of prior knowledge sometimes go along with overconfidence (see Chi, 2006, for a brief review), resulting in reduced effort, self-regulation, and performance in learners (Hall, Ariss, & Todorov, 2007). High prior knowledge in one domain can also hamper learning in a related

domain (negative transfer; Woltz, Gardner, & Bell, 2000). For example, children's highly automatized and correct knowledge about whole numbers can interfere with learning about fractions, which look similar but differ in important mathematical characteristics such as their density (Siegler, Fazio, Bailey, & Zhou, 2013).

The mechanisms described in the context of the knowledge-is-power hypothesis or the knowledge-as-burden hypothesis do not mutually exclude each other. They might work in parallel and interact with each other. This interplay raises questions about the combined net effect of these processes on learning. Is it positive, negative, or close to zero because positive and negative influences cancel each other out?

6.2.2 Relations between Prior Knowledge, Posttest Knowledge and Pretest-Posttest Knowledge Gains

Most studies on the effects of prior knowledge on learning investigated the correlation between prior knowledge and posttest knowledge or the correlation between prior knowledge and pretest-posttest knowledge gains. Knowledge gains can either be computed as absolute gain scores (i.e., *AG = posttest score – pretest score*) or as normalized gain scores (i.e., *NG* = (*posttest score – pretest score*) / (*scale maximum – pretest score*); Hake, 1998). Absolute gain scores have the disadvantage that learners with high prior knowledge have less room on the scale for improvements than learners with low prior knowledge. Normalized gain scores account for this problem by dividing the absolute learning gain by the maximal possible learning gain (Hake, 1998).

Few studies systematically compared the correlations of prior knowledge with posttest knowledge and with (absolute or normalized) knowledge gains. In the present meta-analysis, we argue that the two types of correlations capture different aspects of prior-knowledge effects on learning, are partly independent of each other, and are of different magnitude. Figure 3 shows that both correlations can be highly positive (top left), close to zero (top middle), or highly negative (top right). Highly positive correlation between prior knowledge

and posttest knowledge can also go along with a strongly negative correlation between prior knowledge and normalized knowledge gains (bottom left) or with a correlation close to zero (bottom middle). Figure 3 only visualizes the correlations with normalized knowledge gains r_{NG}, as the graphs for absolute knowledge gains r_{AG} look very similar and normalized knowledge gains have methodological advantages over absolute knowledge gains (Hake, 1998).

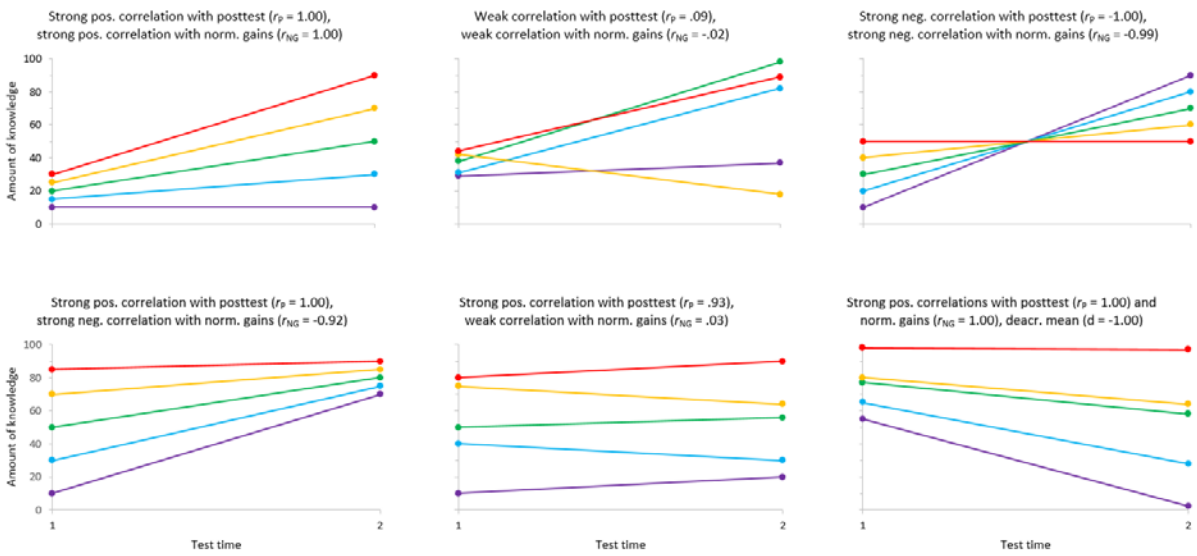

Figure 3. Six different sets of correlations of prior knowledge with posttest knowledge and normalized knowledge gains, visualized using five fictitious persons‘ amount of knowledge before (T1) and after learning (T2) as example.

The correlation r_P between prior knowledge and posttest knowledge indicates to what extent differences between the learners’ amounts of knowledge relative to each other stays stable over time. A strong positive correlation implies that the rank order of the learners with respect to the amount of their knowledge remained relatively unchanged. The correlation r_P can thus be used to predict how well a learner will perform compared to other learners after learning. However, this correlation does not reveal anything about the learners’ absolute amount of knowledge. Strong positive correlations can go along with increases (Figure 3, top

left) as well as decreases (Figure 3, bottom right) in the samples' absolute amount of knowledge over time.

We expected to find a strong positive correlation r_P between prior knowledge and posttest knowledge, because learners' knowledge in a domain typically accumulates over months or years. This is incremental learning is emphasized, for example, by theories of conceptual change or conceptual development (diSessa et al., 2004; Keil, 1996; Vosniadou, 1994), skill building (M. Schwartz & Fischer, 2004), and expertise acquisition (Ullén et al., 2016). Accumulated individual differences that have grown over extended time periods tend to be highly stable and hard to change by short learning interventions (Ericsson, Krampe, & Tesch-Römer, 1993; Vosniadou, Ioannides, Dimitrakopoulou, & Papademetriou, 2001). Thus, we expected that these accumulated and stable individual differences would show in a similar way at pretest as well as posttest, leading to a strong positive correlation.

In contrast, the correlation between prior knowledge and the learning gains indicates to what extent learners with a high amount of prior knowledge have larger knowledge gains ($r > 0$) or smaller knowledge gains ($r < 0$) than their peers. Thus, the correlation can be used to predict to what extent learners with low prior knowledge stay behind, catch up with, or even overtake more knowledgeable learners over time. The correlation does not indicate whether a samples' absolute amount of knowledge increases (Figure 3, top left) or decreases (Figure 3, bottom right) over time, which is indicated by the Cohen's *d* of the pretest-posttest changes.

We expected to find a weak positive correlation r_{NG} between prior knowledge and normalized knowledge gains, because the amount of knowledge gained during a learning phase depends on a multitude of instruction characteristics, teaching characteristics, and learner characteristics (e.g., Ormrod, 2012). Since prior knowledge is just one of these variables, it can only be expected to explain a part of the variance in learners' knowledge gains. We expected that the correlation r_{AG} between prior knowledge and absolute knowledge

gains would be even lower, because absolute gains penalize high prior knowledge and can thus underestimate the true association strength (Hake, 1998).

Either of the two types of correlations alone does not reflect some aspects of the learning process, but only by their combination. For example, when both correlations are strongly positive (Figure 3, top left), the between-learner variance necessarily increases, as learners with already high prior knowledge also have the largest knowledge increases, whereas learners with lower prior knowledge have smaller increases. This phenomenon is also discussed under the label Matthew effect in the literature (e.g., Stanovich, 1986; Walberg & Tsai, 1983). When the correlation between prior knowledge and knowledge gains is negative but the correlation between prior knowledge and posttest knowledge is positive, the variance decreases, because the student with low prior knowledge catch up with their peers (bottom left). When both correlations are negative, the students with low prior knowledge gain so much more knowledge than their peers that the learners' rank order reverses over time (top left). This highlights the importance of analyzing both correlations separately and interpreting them together.

6.2.3 Possible Moderating Influences

Systematic differences between the studies on prior knowledge might explain why some studies found stronger correlations between prior knowledge and posttest knowledge or knowledge gains than others did. In the following, we introduce three groups of possible moderators: (1) knowledge characteristics, (2) learner characteristics, and (3) environmental characteristics.

6.2.3.1 Knowledge Characteristics

A characteristic of knowledge that potentially moderates the relation between prior knowledge and learning is the *type of knowledge*. A widely used distinction between knowledge types is the one between declarative and procedural knowledge (Anderson et al., 2004). Declarative knowledge is verbalizable and explicit. Two educationally relevant

subtypes of declarative knowledge are fact knowledge (e.g. 5 + 2 = 7) (Michael Schneider & Grabner, 2012), which tends to consist of isolated pieces, and conceptual knowledge, which tends to be more abstract and relational (e.g., the principle of commutativity in mathematics) (Goldwater & Schalk, 2016). Procedural knowledge consists of automatized and implicit rules that specify which operators help to reach a goal. It includes knowledge about cognitive operators, for example, how to add two numbers, and motoric procedural knowledge, for example, how to jump (Rittle-Johnson, Schneider, & Star, 2015). A second characteristic of knowledge that might moderate the relation between prior knowledge and learning is the *content domain*. Content domains can differ in how closely and systematically elements of knowledge are inter-related. Prior knowledge might more strongly facilitate learning in domains with higher degrees of interrelatedness than in domains with lower degrees of interrelatedness. A third potential moderator is the *similarity* of prior knowledge and posttest knowledge. The more similar prior knowledge is to new knowledge, the easier it might be for a learner to use the prior knowledge in order to learn the new knowledge (Medin, Goldstone, & Gentner, 1993). This similarity is multidimensional. Among the dimensions on which two pieces of knowledge can be more or less similar are knowledge domain, physical context, temporal context, functional context, social context, and modality of assessment (Barnett & Ceci, 2002).

6.2.3.2 Learner Characteristics

Possible moderators on the level of learner characteristics are age and educational level. *Age* is a potential moderator, because it relates to differences in cognitive processing and learning (e.g., executive functions; Lejeune, Desmottes, Catale, & Meulemans, 2015; Steinberg, 2014). Furthermore, the importance of prior knowledge might differ between learners' *educational levels* (kindergarten, primary school, secondary school, higher education etc.), which vary in learner age, required prior knowledge, and, partly, instructional methods.

6.2.3.3 Environmental Characteristics

Additionally**,** environmental characteristics might moderate the relation between prior knowledge and learning. Students often fail to activate prior knowledge when it would be relevant (Renkl, Mandl, & Gruber, 1996). Thus, instructional interventions designed in ways to activate, explicate, and address prior knowledge likely strengthen the relation between prior knowledge and learning (Baumert et al., 2010; Pressley et al., 1992). For studies in which the learners participated in an instruction intervention, we coded the setting (e.g., school instruction), the duration, the cognitive demands, and the instructional methods employed. We expected a particularly strong moderation effect of cognitive demands. Instruction with high cognitive demands requires learners to understand conceptual connections, analyze, justify, explain, and draw conclusions (Stein & Smith, 1998). This is easier for learners with high prior knowledge than for learners with low prior knowledge (Kirschner et al., 2006). Accordingly, we expected that instruction with high cognitive demands would amplify pre-existing differences in prior knowledge leading to strong positive correlations between prior knowledge, posttest knowledge, and knowledge gains. A related possible environmental moderator is the *country* as countries differ in their educational systems, curricula, and cultural patterns of teaching (Hiebert, Stigler, & Manaster, 1999, p. 196). Many other knowledge, learner, and environmental characteristics proved to be important in single studies, but could not be included in our meta-analysis, because the number of studies was too small.

6.2.4 The Present Study

In sum, the concept of prior knowledge is widely used in many areas of psychological research, but the strength of the effect of domain-specific prior knowledge on learning is still unknown. According to the *knowledge-is-power hypothesis*, the overall effect of prior knowledge is positive; according to the *knowledge-as-burden hypothesis,* it is negative. It is also unclear to what extent the findings obtained with posttest knowledge, absolute

knowledge gains, or normalized knowledge gains differ and how strongly characteristics of the knowledge, the learner, and the environment moderate the effect sizes. We therefore conducted a meta-analysis on these questions. To maximize the generalizability of our findings, we did not restrict our analysis to any content domains, age groups, or institutional contexts. We included all studies using domain-specific prior knowledge to predict knowledge or achievement at posttest or knowledge gains from pretest to posttest. We included only knowledge, and not achievement as an independent variable, because we were interested in the effects of prior knowledge, which is related to but not identical with achievement. We included both knowledge and achievement as dependent variables, because we were interested in the effects of prior knowledge on domain-specific cognitive learning outcomes in general. In the remainder of the manuscript, for simplicity, we used the term *posttest knowledge* to refer to knowledge and achievement at posttest and *knowledge gains* to refer to changes in knowledge from pretest to posttest. Intelligence might artificially inflate the correlation between prior knowledge and learning by affecting the acquisition of prior knowledge and, independently of that, also the acquisition of new knowledge. Therefore, we tested how much controlling for intelligence reduces the correlations.

We had four research questions. Our first research question was how much empirical evidence is available about the relation between prior knowledge and both, posttest knowledge and knowledge gains, and about the causality of these relations. As prior knowledge is a widely-used concept throughout Psychology and Education, we expected to find a large amount of relevant empirical evidence. Our second research question was how prior knowledge relates to posttest knowledge. As described, the correlation between prior knowledge and posttest knowledge indicates the stability of individual differences, which we expected to be high, because these differences can accumulate over months or years. We expected similar findings for knowledge and for achievement as learning outcome measures, because knowledge is a central component of achievement (M. A. Johnson & Lawson, 1998;

O'Reilly & McNamara, 2007; OECD, 2016). Our third research question concerned how prior knowledge predicts pretest-posttest knowledge gains. We expected this correlation to be weakly positive, because prior knowledge is just one of many causal influences on learning. Our fourth research question was how strongly the effects of prior knowledge on posttest knowledge and on knowledge gains are moderated by characteristics of the knowledge, the learner, and the environment. As explained in the introduction, there are reasons to expect that the effect of prior knowledge on learning differs systematically between studies.

6.3 Method

6.3.1 Literature Search

Figure 4 summarizes the literature search process and inclusion criteria. In May 2015, we searched the title, abstract, and keywords of all articles in the literature database PsycINFO. The search string was designed to not only include studies explicitly using the term prior knowledge, but also other studies investigating knowledge in designs with at least two measurement points. The search string was: *(pre-test OR post-test OR pretest OR posttest OR pre test OR post test OR longitudinal OR repeated measure* OR measurement point* OR prior knowledge) AND knowledge*. After limiting the results to empirical quantitative studies conducted with non-disordered participants, written in the English language, and published in a peer-reviewed journal, the search provided 4572 hits (for more details, see Figure 4). We updated the search two years later in May 2017. This search provided additional 890 hits, resulting in 5462 hits in total. We accounted for publication bias by using visual and statistical methods (described in the results section) instead of including unpublished studies. The quality of unpublished studies is hard to assess, and researchers usually obtain only a non-representative set of unpublished studies which does not increase the quality of the meta-analytic results (Ferguson & Brannick, 2011).

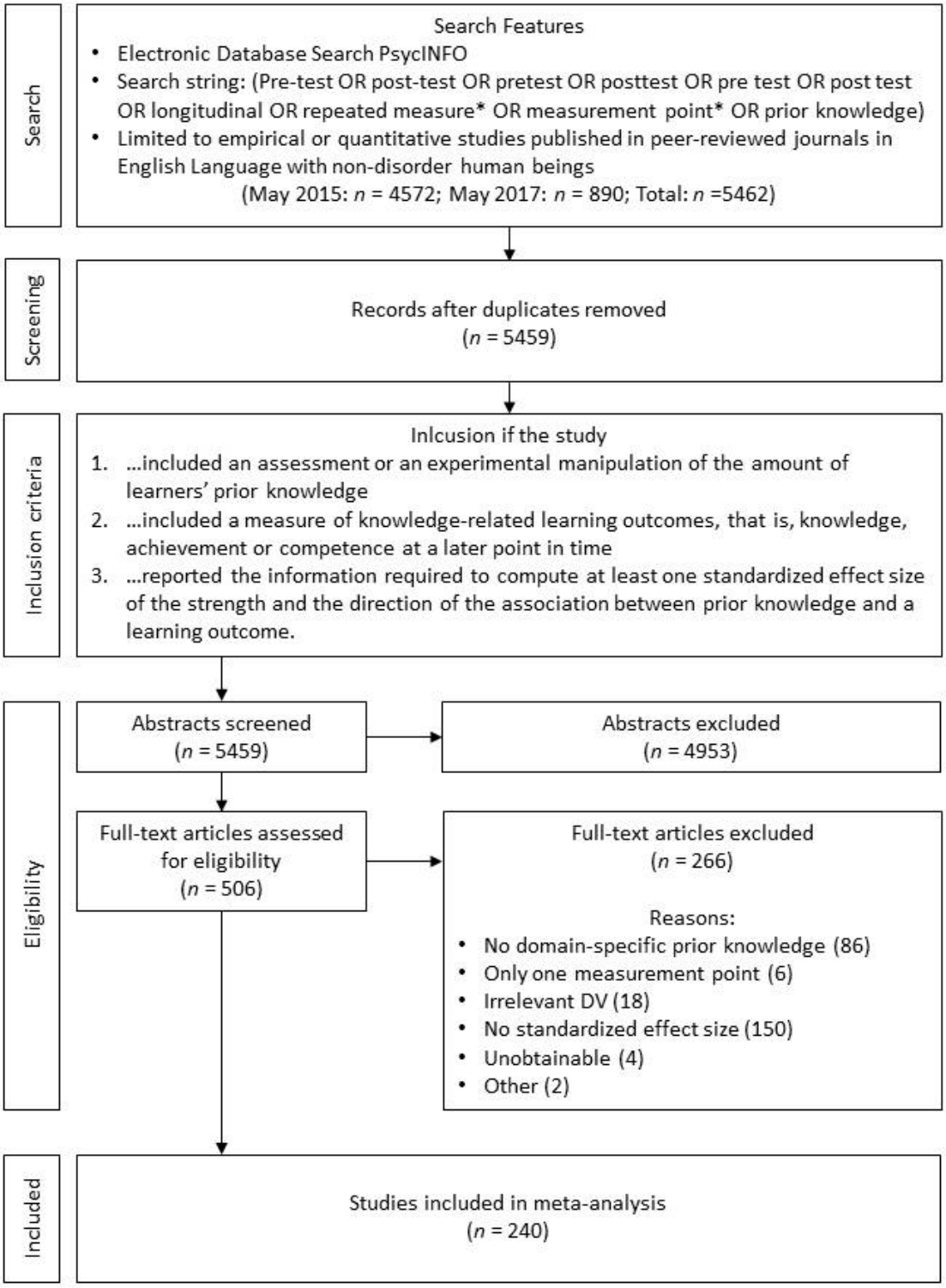

Figure 4. Flow chart for the literature search.

6.3.2 Inclusion of Studies

The three inclusion criteria were: (1) The study included an assessment or an experimental manipulation of the amount of learners' domain-specific prior knowledge as

defined in the introduction. To facilitate the interpretation of the results, we excluded studies in which not the amount of knowledge but the activation of prior knowledge was manipulated (e.g., Amadieu et al., 2015) or which compared learning with familiar versus unfamiliar materials (e.g., Badham, Hay, Foxon, Kaur, & Maylor, 2016). We only included objective quantitative measures of domain-specific prior knowledge, and excluded self-assessments, composite scores from more than one domain, as well as measures of crystallized intelligence, abilities, achievement, or meta-cognitive knowledge. (2) The study included a measure of knowledge or achievement after learning or of the gains made from one measurement point to another. Studies in which the domain of the prior knowledge test and the learning outcome test differed were included. We also included composite measures of learning outcomes in more than one domain (e.g., GPA). We excluded studies reporting learner self-ratings of their knowledge. (3) The study reported the information required to compute at least one standardized effect size index of the strength and the direction of the association between prior knowledge and the outcome after learning or the learning gains.

After removing three duplicates, we screened all remaining 5459 titles and abstracts and excluded studies obviously not relevant for our meta-analysis. For studies with eligible titles and abstracts, we inspected the full texts. The first author screened all studies for inclusion. The second author independently screened a random sample of 115 studies for inclusion. The inter-coder agreement was 83%. Interrater disagreements were resolved through discussion. The studies included in the meta-analysis are listed in Appendix C.

6.3.3 Data Coding

The coding rules were documented in a coding manual, which was used for rater training. The first author coded the information from all included studies. The second author independently coded the information from a random sample of 547 effect sizes from 86 studies. The inter-coder agreement was 90% of all coded values. Inter-coder disagreements

were resolved through discussion. Appendix D provides a list of all coded variables and their levels.

For studies with more than two measurement points and for longitudinal studies, we coded one effect size for each possible pair of measurement points. To aid the interpretation of the results, we coded the levels of the moderator variables only in cases where the levels were the same for pretest and posttest (e.g., educational level was coded only when the students were on the same level at T1 and T2, but not when they were in elementary school at T1 and in middle school at T2). We coded the similarity of prior knowledge and posttest knowledge using the taxonomy proposed by Barnett and Ceci (2002). To ease coding, we only distinguished between two levels of each dimension (similarity high vs. similarity low). For example, we coded the similarity of the temporal context as high when prior knowledge and learning outcomes were assessed on the same day and as low, when they were assessed on different days. In addition to Barnett and Ceci's dimensions, we coded the dimension knowledge type (e.g., high similarity for declarative knowledge at T1 and at T2, low similarity for declarative knowledge at T1 and procedural knowledge at T2). We coded the cognitive demands of the intervention based on the authors' description. We coded this moderator only for studies including at least one condition with higher cognitive demands and one with lower cognitive demands and when the conditions were explicitly described by the authors as having higher or lower cognitive demands relative to each other. When coding the instructional methods, we used the labels employed by the authors as our main criterion. For example, we only coded the employment of problem-based learning, when the authors of the original publication used the term problem-based learning to describe their instructional intervention.

6.3.4 Preparation of Effect Sizes

Appendix E lists the equations for the preparation of the effect sizes for the meta-analytic integration. We used Pearson correlations (r) as the dependent variable of our meta-analysis,

because the majority of the included studies reported correlational results. Whenever a study compared a high- and a low-prior knowledge group and reported the group means and standard deviations for at least two measurement points, we calculated the absolute and the normalized gain scores from this information. Absolute learning gains were computed as *posttest score – pretest score*, and normalized learning gains were computed as (*posttest score – pretest score*) / (*scale maximum – pretest score*) (Hake, 1998). We then computed group differences in these gain scores as Cohen's *d* values, respectively. For studies reporting group differences for the posttest only, we computed Cohen's *d* for this difference. We then converted the Cohen's *d*s to Pearson correlations. Unless specified otherwise, all correlations were coded so that a positive value indicated that a higher quantity of (correct or incorrect) prior knowledge went along with a higher quantity of (correct or incorrect) posttest knowledge or positive pretest-posttest gains. The correlations were subjected to a Fisher's Z_r-transformation to approximate a normal sampling distribution (Lipsey & Wilson, 2001).

Only when original studies reported the reliabilities of the measures, we corrected the correlations for measurement error, which would otherwise make the correlation between two constructs appear smaller than it actually is (F. L. Schmidt & Hunter, 2015, p. 112). Some studies measured prior knowledge as a continuous variable and then dichotomized the participants into a high- and a low-prior knowledge group. We corrected for this loss of information (i.e., loss of variance) using the formula given by F. L. Schmidt and Hunter (2015, p. 134). We winsorized the corrected correlations and variances to less extreme values when they exceeded a value above -1 or 1 due to attenuation.

6.3.5 Statistical Analysis

6.3.5.1 Outliers

Before we performed the meta-analysis, we identified outliers through Cook's values (Cook & Weisberg, 1982; Viechtbauer & Cheung, 2010) using the *metafor* package (Viechtbauer, 2010) in R. We excluded two outliers found with posttest knowledge (Lonigan,

Burgees, & Anthony, 2000; Ree, Carretta, & Teachout, 1995), two found with absolute knowledge gains (Zacharia, Loizou, & Papaevripidou, 2012), and one found with normalized knowledge gains (Zacharia et al., 2012).

6.3.5.2 Publication Bias

We visually and statistically tested for publication bias using funnel plots and Egger regressions for random-effects models (Egger et al., 1997). These tests were conducted using the *metafor* package (Viechtbauer, 2010) in R (R Core Team, 2014).

6.3.5.3 Meta-Analytic Integration

The formulas used for the meta-analytic integration of the effect sizes are provided in Appendix E. The majority of the included studies reported several effect sizes, for example, for various dependent measures or measurement points. These effect sizes are statistically dependent and thus violate a central assumption of classical meta-analytical models. To handle statistically dependent effect sizes, we employed robust variance estimation (Hedges et al., 2010; Tanner-Smith et al., 2016). Given the expected heterogeneity, we used random effects models for the meta-analytic integration of the effect sizes (cf. Raudenbush, 2009). Mean effect sizes and meta-regression models were estimated using a weighted least squares approach (cf. Hedges et al., 2010; Tanner-Smith & Tipton, 2014). The statistical analyses were performed using the *robumeta* package (Z. Fisher & Tipton, 2014) in R.

6.3.5.4 Moderators

We computed the mean effect size for every level within a moderator category separately. For levels with degrees of freedom smaller 4, we only report the mean, but not the confidence interval, as results are not trustworthy due to the small number of observations. These levels were also excluded from the significance tests of the moderators. Before performing the moderator analyses, continuous moderator variables were log-transformed to obtain normal distributions of the variables. Categorical moderators were dummy-coded using the moderator level with the lowest effect size as the reference level. Unless stated otherwise, each

moderator was tested in a separate analysis. The moderators were entered as predictors in regression models for the prediction of the effect sizes (see Appendix E). For dummy coded variables, each predictor indicates whether the coded moderator level significantly differs from the reference level. We computed the overall significance for each regression analysis using the Wald-test function of the *clubSandwich* package in R (Pustejovsky, 2017; Tanner-Smith et al., 2016). We computed the overall proportion of explained variance R^2 for each regression as described in Appendix E.

6.4 Results

6.4.1 Characteristics of the Included Studies

The 240 included articles reported results from 335 independent samples with 4327 relevant effect sizes and 62,129 participants in total. The publication years ranged from 1965 to 2017, with a median of 2010. Of the 240 articles, 60% were published within the last 10 years (2008-2017), indicating an increasing interest in prior knowledge. Appendix D lists the frequencies of the levels of the moderator variables in the included studies. Twenty-eight countries are represented in the meta-analysis. Most studies reported data from North America (45%), followed by Europe and the Middle East (32%), Asia (11%), and Australia/New Zealand (1%). The time between the measurement of prior knowledge (T1) and the measurement of the learning outcomes (T2) varied between 0 and 3780 days, with a median of 360 days. Regarding the learners' age, sample means ranged from 0.63 to 33.25 years with an overall mean of 11.13 years (SD = 7.02).

6.4.2 Characteristics of the Included Effect Sizes

Our first research question was how much empirical evidence is available on the correlational and causal relations of prior knowledge with posttest knowledge and knowledge gains, respectively. We found a large number of studies that investigated the correlation between prior knowledge and posttest knowledge. This relationship had been investigated in

235 included studies, reporting 4223 effect sizes ranging between r_p^+ = -.635 to r_p^+ = .995. Of these, 44 effect sizes from nine studies were obtained in randomized controlled trials (RCTs). We categorized a study as RCT when the participants were randomized into at least two groups and the levels of prior knowledge were manipulated to differ between these groups. For 28 studies with 1305 effect sizes, it was possible to control for intelligence.

In contrast, the association between prior knowledge and knowledge gains had been investigated in only few studies, none of which was an RCT or allowed controlling for intelligence. The correlation between prior knowledge and absolute learning gains could be computed for 20 studies with 62 effect sizes ranging from r_{AG}^+ = -.971 to r_{AG}^+ = .726. The correlation between prior knowledge and normalized learning gains could be computed for 14 studies with 33 effect sizes ranging from r_{NG}^+ = -.952 to r_{NG}^+ = .966. Thus, overall, there is a strong asymmetry in how much empirical evidence is available about prior-knowledge effects on posttest knowledge and prior-knowledge effects on knowledge gains.

There was no evidence for a publication bias, that is, an underrepresentation of effect sizes close to zero, in our database. The effect sizes found with posttest knowledge, absolute knowledge gains, and normalized knowledge gains varied about symmetrically around their common means, both for the individual effect sizes (Figure 5, left column) and for the study-average effect sizes (Figure 5, right column). Neither Egger regressions or the trim-and-fill method (Duval & Tweedie, 2000) indicated any underrepresentation of small effect sizes in any of the six cases depicted in Figure 5.

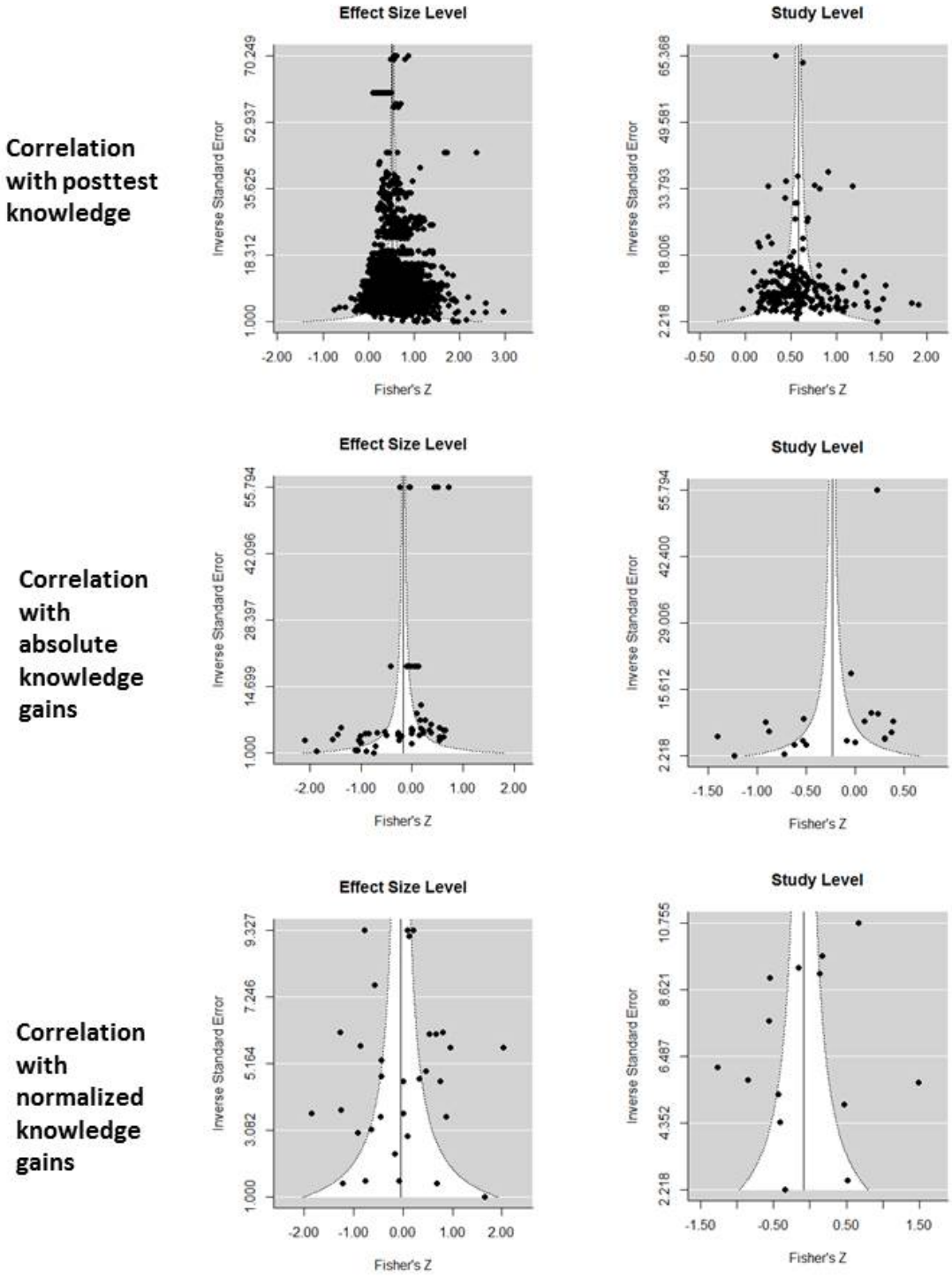

Figure 5. Funnel plots of the inverse of standard error and the effect size (corrected for artifacts and here transformed to Fisher's Z) for the effect-size level (left) and the study level of aggregation (right) for learning outcomes (top) and learning gains (bottom).

6.4.3 Correlation between Prior Knowledge and Posttest Knowledge

Table 3 shows the average effect sizes separately for studies using prior knowledge to predict posttest knowledge, absolute knowledge gains, or normalized knowledge gains. With respect to our second research question, i.e., how strongly prior knowledge is associated with posttest knowledge, we found a strong, positive, and statistically significant correlation of r_P^+ = .525, as expected. The correlation is strong by the standards of Cohen (1992), indicating a high stability of individual differences in knowledge from before to after learning. The I^2 index of 93.53 indicated a large amount of heterogeneity, implying that third variables moderate the extent of this stability.

The connection between prior knowledge and posttest knowledge is causal, as indicated by the significant positive effect size of r_P^+ = .308 found in RCTs. The correlation was significantly lower in RCTs than in other study designs, but this difference explained only a variance proportion of R^2 = .007 of the effect sizes. The effect of intelligence on prior knowledge and learning outcomes contributed only weakly to the high correlation. Controlling the correlation for intelligence across the 28 relevant studies for which this was possible, the correlation only marginally changed from r_P^+ = .518, 95% CI [.464, .568], I^2 = 90.31, to r_P^+ = .482, 95% CI [.424, .535], I^2 = 91.75. The difference between these two correlations was not statistically significant (z = 1.23, p = .219). Therefore, the association of prior knowledge with posttest knowledge is causal and cannot entirely be attributed to a confounding influence of intelligence.

6.4.4 Correlation between Prior Knowledge and Knowledge Gains

With respect to our third research question, i.e., how strongly prior knowledge is associated with knowledge gains, we expected to find weak positive effects. Against our expectation, the empirical findings were inconclusive. The correlation was descriptively negative but not significantly smaller than zero for absolute knowledge gains (r_{AG}^+ = -.210) as well as for normalized knowledge gains (r_{NG}^+ = -.083). The 95% confidence intervals around

these means are extremely large (see Table 3), because only 20 studies reported the correlation between prior knowledge and absolute knowledge gains (or the information required to compute this correlation) and a mere 14 studies reported the correlation between prior knowledge and normalized knowledge gains (or the information required to compute this correlation). The heterogeneity indices I^2 were close to 100%, indicating high degrees of heterogeneity of the effect sizes found with absolute knowledge gains and with normalized knowledge gains. Due to a lack of relevant empirical studies, the causality of the prior knowledge effects on knowledge gains as well as the influence of intelligence could not be investigated in our meta-analysis.

6.4.5 Differences between the Correlations Found with the Three Types of Dependent Variables

As expected, the correlation was highest for posttest knowledge as dependent variable (r_p), lower for normalized knowledge gains (r_{NG}), and lowest for absolute knowledge gains (r_{AG}). The confidence interval for posttest knowledge did not overlap with the confidence intervals for absolute and normalized knowledge gains, which indicates a statistically reliable difference. The difference between posttest knowledge and knowledge gains was significant (z = 6.07, p = <.001 for absolute knowledge gains and z = 3.64, p = <.001 for normalized knowledge gains). In contrast, the confidence intervals for absolute and normalized knowledge gains overlapped strongly, and their descriptive difference did not reach statistical significance (z = 0.58, p = .562).

We repeated the analyses for those 12 studies that reported sufficient information only, to compute and compare all three types of dependent variables (thus holding study characteristics constant over the three types of dependent variables; see Table 3). This increased the differences and led to mean effect sizes ranging from r_P^+ = .760 for posttest knowledge over r_{NG}^+ = .123 for normalized knowledge gains, to r_{AG}^+ = -.398 for absolute knowledge gains. These results indicate that, independently of other study characteristics,

correlations with posttest knowledge and with knowledge gains are significantly different (all *p*s < .05), capture different and partly independent aspects of the effect of prior knowledge on learning, and need to be analyzed separately. In the following two sections, we thus report the moderator analyses separately for posttest knowledge and knowledge gains.

6.4.6 Moderators of the Correlation between Prior Knowledge and Posttest Knowledge

Our fourth research question considered how strongly the effect of prior knowledge on posttest knowledge and on knowledge gains is moderated by knowledge-related, learner-related, and environment-related variables. Table 4 shows the overall mean correlation between prior knowledge and posttest knowledge and how this overall effect was moderated by third variables. When levels of moderators were coded (see Appendix D) but are not listed in Table 4, maximally one study reported the respective level of the moderator. Likewise, if confidence intervals or tests results of moderator effects are not listed in Table 4 indicates, that the available evidence was too limited to permit these analyses. Moderator levels for which the evidence was so limited that no confidence intervals could be computed were excluded from the significance tests.

Table 3.

Meta-analytic overall results separately for the three dependent variables learning outcome, absolute learning gains, and normalized learning gains

	Correlation with posttest knowledge						Correlation with absolute knowledge gains						Correlation with normalized knowledge gains					
	j	*k*	r_P^+	CI r_P^+ 95%	τ^2	I^2	*j*	*k*	r_{AG}^+	CI r_{AG}^+ 95%	τ^2	I^2	*j*	*k*	r_{NG}^+	CI r_{NG}^+ 95%	τ^2	I^2
Studies allowing the computation of at least one of the three types of effect sizes	235	4223	.525	[.495, .553]	.099	93.53	20	62	-.210	[-.437, .043]	.380	96.99	14	33	-.083	[-.478, .342]	.580	94.83
Studies allowing the computation of all three types of effect sizes	12	30	.760	[.603, .864]	.362	90.09	12	30	-.398	[-.713, .050]	1.20	96.78	12	30	.123	[-.352, .556]	1.05	96.34
Randomized controlled trial																		
No	226	4179	.532	[.502, .561]	0.099	93.70	20	62	-.210	[-.437, .043]	.380	96.99	14	33	-.083	[-.478, .342]	.580	94.83
Yes	9	44	.308	[.186, .420]	0.038	60.61	0	0	-	-	-	-	0	0	-	-	-	-
Controlling for intelligence																		
Before controlling	28	1305	.518	[0.464, 0.568]	0.061	90.31	0	0	-	-	-	-	0	0	-	-	-	-
After controlling	28	1305	.482	[0.424, 0.535]	0.072	91.75	0	0	-	-	-	-	0	0	-	-	-	-

Table 4.

Mean effect sizes and moderator analyses examining learning outcomes.

	j	*k*	r_P^+	CI r_P^+ 95%	τ^2	I^2	Moderator	
							Sign.	R^2
Knowledge characteristics								
Knowledge type[a]							*	.037
Declarative	159	2125	.513	[.478, .548]	.205	95.68	Ref	
Procedural	10	51	.645	[.468, .774]	.117	91.83	*ns*	
Declarative and procedural mixed	74	645	.619	[.566, .667]	.104	95.51	**	
Knowledge sub-type[a]							*ns*	.016
Declarative: facts	41	268	.580	[.507, .645]	.153	94.28	*ns*	
Declarative: concepts	128	1244	.528	[.487, .566]	.220	96.21	Ref	
Procedural: cognitive skill	10	51	.645	[.468, .774]	.117	91.83	*ns*	
Broad content area[a]							*ns*	.002
STEM	87	578	.566	[.509, .619]	.112	943.81	*ns*	
Language	99	3055	.540	[.504, .575]	.072	92.11	Ref	
Humanities	5	19	.355	-	-	-	-	
Social sciences	17	107	.570	[.390, .708]	.197	93.69	*ns*	
Health sciences	2	9	.439	-	-	-	-	
Sports	5	54	.533	-	-	-	-	
Content domain[a]							*ns*	.022
Mathematics	37	416	.601	[.521, .670]	.087	94.23	**	
Physics	16	55	.546	[.418, .653]	.139	93.85	*	
Chemistry	6	20	.352	[.149, .528]	.054	97.42	Ref	
Biology	21	62	.624	[.474, .739]	.250	91.51	**	
Geosciences	6	19	.482	[.002, .782]	.245	92.83	*ns*	
Medicine and nursing	2	9	.439	-	-	-	-	
Psychology	14	99	.546	[.401, .665]	.150	92.41	*	
First language	94	2715	.537	[.499, .573]	.076	92.20	**	
Second language	10	53	.710	[.610, .789]	.076	87.52	**	
Sports	5	54	.533	-	-	-	-	

History	3	10	.457	-	-	-	-	
Other	3	20	.503	-	-	-	-	
Similarity of prior knowledge and learning outcome								
Similarity of the knowledge type							**	.031
Low	117	1398	.441	[.409, .472]	.053	91.39	Ref	
High	210	2821	.547	[.515, .578]	.162	95.63	**	
Similarity of the content domain							**	.036
Low	72	688	.419	[.364, .472]	.044	89.82	Ref	
High	209	3535	.548	[.518, .578]	.103	94.01	**	
Similarity of the physical context							*ns*	.000
Low	7	94	.566	[.373, .712]	.230	96.40	*ns*	
High	228	4071	.528	[.498, .558]	.134	95.30	Ref	
Similarity of the temporal context							*ns*	.000
Low	184	3974	.529	[.496, .560]	.096	94.16	*ns*	
High	55	248	.515	[.441, .581]	.129	88.13	Ref	
Time between assessments	171	3908					*ns*	.000
Similarity of the social context							*	.008
Low	15	185	.451	[.357, .535]	.037	88.43	Ref	
High	222	3909	.533	[.502, .561]	.112	93.30	**	
Similarity of the modality							**	.026
Low	46	650	.437	[.381, .490]	.084	92.86	Ref	
High	205	3064	.538	[.505, .570]	.155	95.56	**	
Multivariate effect of all similarity dimensions	204	3397					*	.021
Learner characteristics								
Age	136	3362					*ns*	.000
Educational level							**	.059
Daycare	5	139	.377	-	-	-	-	

Kindergarten/preschool	57	1790	.458	[.411, .504]	.060	90.16	*ns*	
Primary education	71	1676	.584	[.542, .624]	.064	93.80	**	
Secondary education	49	228	.543	[.455, .621]	.125	92.87	*ns*	
Higher education	57	358	.457	[.392, .518]	.108	90.46	Ref	
Continued education	5	22	.753	-	-	-	-	
Several	3	10	.743	-	-	-	-	
Environmental characteristics								
Intervention								
Intervention Setting							*ns*	.001
No intervention	14	150	.487	[.356, .599]	.073	83.16	Ref	
School instruction only	78	2772	.511	[.461, .557]	.106	96.17	*ns*	
School instruction and other intervention	57	867	.513	[.451, .570]	.075	92.49	*ns*	
Other intervention only	85	397	.548	[.489, .602]	.145	89.04	*ns*	
Intervention Duration							*ns*	.008
0-2 hours	73	352	.515	[.452, .573]	.141	88.93	*ns*	
2-24 hours	7	29	.497	[.197, .711]	.152	83.31	Ref	
2-7 days	5	16	.516	-	-	-	-	
> 1 week	57	889	.561	[.496, .621]	.086	93.60	*ns*	
Cognitive demands of intervention							*	.056
Lower	39	100	.447	[.329; .551]	.162	94.42	Ref	
Higher	38	96	.656	[.543; .745]	.281	90.06	**	
Instructional methods in intervention								
Written instruction							*ns*	.007
No	23	226	.571	[.464, .662]	.171	90.93	*ns*	
Yes	101	463	.519	[.469, .565]	.101	87.04	Ref	
Oral instruction							*ns*	.016
No	85	495	.504	[.452, .553]	.106	85.98	Ref	
Yes	36	253	.584	[.495, .660]	.163	92.23	*ns*	
Multimedia instruction							*ns*	.012

No	70	489	.496	[.438, .550]	.098	85.44	Ref	
Yes	43	149	.585	[.508, .653]	.138	87.66	*ns*	
Practice							*ns*	.006
No	75	440	.516	[.457, .570]	.114	87.43	Ref	
Yes	47	302	.563	[.485, .633]	.147	91.91	*ns*	
Constructive activities							*ns*	.017
No	81	501	.504	[.451, .555]	.098	83.64	Ref	
Yes	45	158	.608	[.516, .685]	.207	92.22	*ns*	
Technology							*ns*	.018
No	78	521	.505	[.449, .557]	.102	86.72	Ref	
Yes	41	127	.630	[.525, .717]	.223	91.94	*ns*	
Feedback							*ns*	.024
No	94	561	.520	[.469, .569]	.124	87.71	Ref	
Yes	18	65	.640	[.476, .761]	.207	91.74	*ns*	
Collaborative learning							*ns*	.020
No	98	569	.514	[.466, .559]	.104	87.16	Ref	
Yes	22	87	.627	[.478, .742]	.228	92.65	*ns*	
Problem-based learning							*	.012
No	96	583	.550	[.495, .600]	.132	88.22	*	
Yes	17	93	.451	[.371, .524]	.073	87.03	Ref	
Multivariate effect of all instructional methods	98	582					*ns*	.088
Country[b]							*ns*	.088
Australia	3	29	.375	-	-	-	-	
Belgium	2	80	.231	-	-	-	-	
Canada	5	115	.715	[.526, .837]	.142	91.90	**	
Peoples Republic of China	5	300	.482	[.401, .556]	.108	93.55	**	
Hong Kong	6	168	.380	[.274, .477]	.071	92.80	Ref	
Taiwan	3	9	.675	-	-	-		
Denmark	2	7	.371	-	-	-		
United Kingdom	12	627	.508	[.408, .595]	.074	89.96	*	
Finland	8	322	.425	[.262, .566]	.088	92.45	*ns*	

France	2	122	.333	-	-	-	-	
Germany	9	253	.491	[.223, .690]	.141	97.91	*ns*	
Israel	4	49	.471	-	-	-	-	
Netherlands	5	163	.643	[.180, .872]	.216	96.45	*ns*	
New Zealand	2	42	.814	-	-	-	-	
Norway	3	172	.557	-	-	-	-	
Spain	2	5	.467	-	-	-	-	
USA	52	904	.518	[.450, .580]	.123	96.45	**	
Methodological study characteristics								
Study design							*ns*	.002
Group differences in prior knowledge (quasi-experimental or experimental design)	83	403	.549	[.488, .604]	.140	88.12	*ns*	
Individual differences in prior knowledge (correlational design)	152	3820	.514	[.480, .547]	.093	94.84	Ref	
Number of items in prior knowledge measure	186	3053					**	.036
Number of items in learning outcome measure	189	3088					*	.025
Response format[a]							*ns*	.019
Open	101	1368	.521	[.472, .567]	.124	92.85	Ref	
Fill-in	7	34	.632	[.562, .693]	.015	58.05	*	
Single or multiple choice	66	508	.556	[.501, .606]	.096	90.52	*ns*	
Other	6	30	.528	[.312, .692]	.117	94.85	*ns*	
Various	18	99	.619	[.506, .711]	.166	95.51	*	
Retention test							*ns*	.000
No	235	4159	.524	[.495, .553]	.099	93.57	*ns*	
Yes	13	64	.507	[.366, .626]	.077	83.03	Ref	
Same test at T1 and T2							**	.149
No	201	3606	.473	[.444, .501]	.070	91.27	Ref	
Yes	104	615	.681	[.644, .714]	.180	96.68	**	
Measures at T2								
Outcome at T2							*ns*	.002

Knowledge	228	4169	.528	[.498, .556]	.079	91.63		
Achievement	10	54	.438	[.209, .621]	.352	99.14		
Domain specificity							*ns*	.007
Specific domain	232	4179	.520	[.491, .548]	.081	92.11		
Various domains	8	44	.591	[.171, .831]	.302	98.95		

[a] Only for effects in which the moderator had the same level at T1 and T2. For effects with differing moderator levels see the results category "Similarity of prior knowledge and learning outcome".

[b] Only studies including any kind of school instructions are presented.

Note. * $p < .05$, ** $p < .01$, – *df* was too small for valid interpretation, we therefore do not report the confidence interval and heterogeneity; *ns* = nonsignificant. Moderation analysis was calculated for moderating variables with full rows only, respectively.

6.4.6.1 Knowledge-Related Moderators

Among the knowledge-related moderators, we found significant moderator effects for the knowledge type, which explained a variance proportion of $R^2 = .037$ of the correlation between prior knowledge and posttest knowledge. No moderator effects were found for the knowledge sub-types, the broad content area, or the content domains.

The similarity of the prior knowledge measure and the posttest measure significantly moderated the correlation between prior knowledge and posttest knowledge. Separate bi-variate regressions showed that the correlation was significantly related to the similarity on the dimensions knowledge type ($R^2 = .031$), content domain ($R^2 = .036$), social context of the assessment ($R^2 = .008$), and modality of the assessment ($R^2 = .026$). In all of these cases, the correlation was stronger for a high similarity and lower, but still significantly greater than zero, for a low similarity. When all similarity dimensions were simultaneously entered as predictors into a multiple regression, they explained a variance proportion of $R^2 = .021$. This

relatively small value indicates that participants can still make use of prior knowledge when prior knowledge and new knowledge are low in similarity.

6.4.6.2 Learner-Related Moderators

The participants' age did not moderate the correlations, but educational level was a significant moderator with R^2 = .059. The association between prior knowledge and posttest knowledge was significantly stronger for primary education as compared to higher education.

6.4.6.3 Environment-Related Moderators

We did not find statistically significant moderating effects of country, intervention, and intervention duration. In contrast, the cognitive demands of the interventions significantly moderated the correlation between prior knowledge and posttest knowledge with R^2 = .056. The correlation was lower for interventions with lower cognitive demands and higher for interventions with higher cognitive demands. Only one of the nine instructional methods listed in Table 4 moderated the correlation. Specifically, interventions including problem-based learning had a smaller correlation than other interventions (R^2 = .012). When we entered all instructional methods together as predictors in a multiple regression, no significant effect was found.

6.4.6.4 Methodological Moderators

The correlation between prior knowledge and posttest knowledge was independent of whether group differences or individual differences in prior knowledge were used to predict the learning outcomes. The correlation increased with the numbers of items used in the prior knowledge test (R^2 = .036) and the posttest (R^2 = .025). A likely explanation for this effect is that longer tests typically increase the precision of measurement by reducing the random noise in the data. The item response format, and whether the learning outcome test was a posttest (directly after the intervention) or a retention test (some time after the posttest) did not moderate the correlation. The correlation of prior knowledge and posttest knowledge were significantly higher when the exact same test was used at both measurement points than when

different tests were used, R^2 = .149. Whether prior knowledge was used to predict a knowledge or achievement measure and knowledge in a domain or multiple domains at posttest did not make a statistically significant difference.

6.4.6.5 Differences Between Correct and Incorrect Knowledge

In the analyses reported in Tables 3 and 4, we tested how the amount of prior knowledge predicted the amount of subsequent knowledge and achievement, irrespective of the correctness of knowledge. Here, we tested whether the correctness of knowledge (i.e., misconceptions and error scores vs. correct concepts and solution rates) influenced the direction of the correlation. We gave all scores quantifying the amount of incorrect knowledge a negative sign and all scores quantifying the extent of correct knowledge a positive sign, and recoded the effect sizes accordingly. Table 5 displays the results of the analyses. As expected, the amount of correct prior knowledge correlated positively with the amount of correct posttest knowledge and negatively with the amount of incorrect posttest knowledge. The amount of incorrect prior knowledge correlated positively with the amount of incorrect posttest knowledge and negatively with the amount of correct posttest knowledge. Thus, correct prior knowledge is directly and incorrect knowledge is inversely related to positive performance at posttest. The correctness of knowledge explained a variance proportion of R^2 = .190 of the effect sizes. We repeated the analysis with the absolute values of the effect sizes, thus ignoring the signs and comparing only the strength of the correlations. This analysis did not indicate significant differences between the four correlations. Thus, the correlations of correct and incorrect prior knowledge with learning outcomes differ in their directions, but not in their absolute strengths.

Table 5.

Meta-analytic results for the different combination of correctness of knowledge at T1 and T2

	j	*k*	r_P^+	CI r_P^+ 95%	τ^2	I^2	Moderator	
							Sign.	R^2
Overall	228	4149	.507	[.475, .538]	.128	95.00		
Measures							**	.190
T1: amount of correct knowledge, T2: amount of correct knowledge	225	4037	.524	[.494, .553]	.118	94.63	**	
T1: amount of correct knowledge, T2: amount of incorrect knowledge	9	82	-.557	[-.746, -.287]	.156	90.98	Ref	
T1: amount of incorrect knowledge, T2: amount of correct knowledge	6	18	-.457	[-.587, -.303]	.023	66.22	**	
T1: amount of incorrect knowledge, T2: amount of incorrect knowledge	4	12	.578	-	-	-		

Note. Effect sizes were recoded so that they had a positive sign when correct knowledge directly predicted correct knowledge or when incorrect knowledge directly predicted incorrect knowledge (and a negative sign when this was not the case).

6.4.7 Moderators of the Correlation between Prior Knowledge and Knowledge Gains

We had planned to conduct the same moderator analyses for absolute and normalized knowledge gains as we conducted for posttest knowledge. However, only very few studies reported the information required to compute the correlation between prior knowledge and absolute or normalized knowledge gains. The results of these studies had a high heterogeneity. Thus, only few moderator analyses were possible, and, for these, the statistical power was low. Accordingly, the estimated confidence intervals were very large, and almost no significant moderator effects were found. Appendix D lists the mean values for the levels

of the moderators that could be coded, and Appendix F lists the results of the significance tests that were possible despite the limited amount of data. The only moderator that had significant effects on the correlations with both, absolute and with normalized gain scores, was the cognitive demands of the intervention. The correlation was higher for interventions with high cognitive demands and lower for interventions with low cognitive demands. This difference explained variance proportions of R^2 = .281 for absolute knowledge gains, and of R^2 = .366 for normalized knowledge gains. These moderator effects were very strong, which explains why the effect sizes found with knowledge gains were so heterogeneous and close to zero, and why the results of other moderator analyses not accounting for the cognitive demands of the intervention did not reach statistical significance. Another moderator was only significant for absolute knowledge gains, but not normalized knowledge gains. The correlation was higher for studies using a different tests to assess prior knowledge and learning outcomes compared to studies using the same test before and after learning (R^2 = .110).

6.5 Discussion

In this meta-analysis, we estimated the effect of the amount of prior knowledge on learning. We first discuss the main findings with respect to the four research questions, followed by the limitations of our study and the implications of our findings.

6.5.1 Main Findings

6.5.1.1 Available Empirical Evidence

Our first research question was how much empirical evidence is available on the relation between prior knowledge and learning. We found a strong imbalance between a high number of studies investigating the correlation between prior knowledge and posttest knowledge and a low number of studies investigating the correlation between prior knowledge and knowledge gains. Furthermore, none of the few studies reporting knowledge gains were

randomized controlled trials, and none of these studies allowed controlling for intelligence. Due to the small number of studies on knowledge gains and their high heterogeneity, the main effects and some moderator effects could only be estimated with low precision, and many moderator effects could not be tested at all. In sum, very little is known about how prior knowledge causally affects the learners' increases in knowledge relative to each other over time.

There are at least four explanations for the scarcity of studies on knowledge gains. One contributing factor is a relatively large group of studies in our meta-analysis that had the main aim of comparing the effectivity of instructional conditions. In these studies, knowledge was only used as a covariate. A second contributing factor is that the included longitudinal studies frequently used different knowledge tests at different time points in order to keep the tests age-appropriate, so that no gain scores could be computed. A third contributing factor are studies reporting multivariate analyses of knowledge changes over time and failing to report the zero-order correlations between prior knowledge and subsequent knowledge that are required for the meta-analytic aggregation of the data. Path coefficients from multivariate analyses cannot easily be meta-analyzed, because their values and meanings depend on all other variables in the model, so that the average effect size is difficult to interpret. A fourth contributing factor is that gain scores were once thought to have undesirable statistical characteristics, for example, chronically low reliabilities. Newer studies of gains scores (e.g., Zimmermann & Williams, 1998) found these conclusions to be incorrect and based on unrealistic statistical assumptions (e.g., the same variance at pretest and posttest, which is often not the case, as shown in Figure 3). These newer and more advanced studies showed that gain scores usually have acceptable validities and reliabilities (May & Hittner, 2010) and are useful tools for investigating change (Maris, 1998).

We cannot exclude the possibility that some of the studies investigating knowledge gains were not found in our literature search. However, we assume that the number of these studies

is small, because we used a very broad search string that was designed to include all studies measuring knowledge at more than one measurement point and which was not limited to find studies using the term “prior knowledge”. The search string yielded 5462 studies, which we scanned for inclusion. Only 20 of these studies reported or allowed us to compute the correlation of prior knowledge with absolute learning gains and 14 of these studies reported or allowed us to compute the correlation of prior knowledge with normalized learning gains. Therefore, the low number of studies on knowledge gains is unlikely to be an artifact created by our literature search.

6.5.1.2 Correlations of Prior Knowledge with Posttest knowledge and Knowledge Gains

Our second research question was how strongly prior knowledge is associated with posttest knowledge. Averaged over 4223 effect sizes from 235 studies, we found a correlation of $r_P^+ = .525$, which indicated a high stability of the individual differences in the amount of knowledge over time. Accordingly, the learners’ rank order with respect to the amount of their knowledge tends to remain relatively unchanged between measurement points. The estimation had a high precision as indicated by the small confidence interval ([.478, .548]).

The correlation decreased only slightly and remained statistically significant when only analyzing the nine randomized controlled trials in our database. These nine studies randomized the participants into two groups and induced prior-knowledge differences between these groups, thus indicating a causal connection between prior knowledge and posttest knowledge. Further, the correlation between prior knowledge and posttest knowledge decreased only marginally when controlling for intelligence. Since intelligence influences the acquisition of prior knowledge as well as the acquisition of new knowledge (Schneider & McGrew, 2012, p. 123), it is surprising that controlling for intelligence did not reduce the effect sizes more strongly. One possible explanation is that many studies that included an intelligence measure were studies on language development that measured non-verbal intelligence. Due to this mismatch of verbal knowledge assessments and non-verbal

intelligence measures, the included studies might have underestimated the true correlation between knowledge and intelligence. Further research on this question is needed.

Our finding of a high stability of individual differences in knowledge matches the findings from studies on other cognitive learning outcomes, namely expert performance and academic achievement. Like individual differences in knowledge, individual differences in expert performance have been found to be highly stable over time (Ericsson & Lehmann, 1996). Previous meta-analyses on achievement (Bretz Jr., 1989; P. A. Cohen, 1984; G. J. Duncan et al., 2007; Kuncel, Hezlett, & Ones, 2001; La Paro & Pianta, 2000; Samson, Graue, Weinstein, & Wahlberg, 1984; Schuler, Funke, & Baron-Boldt, 1990; Trapmann, Hell, Weigand, & Schuler, 2007) found correlations of prior achievement with later achievement ranging from r = .10 (G. J. Duncan et al., 2007) to r = .48 (La Paro & Pianta, 2000). Our average correlation found for knowledge lies slightly above this range, possibly because knowledge is a less heterogeneous concept than achievement and is thus easier to predict over time. A second explanation is that achievement is mostly measured in longitudinal studies with month or years between measurement points, whereas prior knowledge is also investigated in smaller intervention studies where pretest and posttest were just hours or days apart (e.g., Fyfe & Rittle-Johnson, 2016)

Our third research question was how strongly prior knowledge correlates with absolute and normalized knowledge gains. We expected to find weak positive effect sizes, because many variables causally determine learning, with prior knowledge being just one of them (Ormrod, 2012). Indeed, the meta-analytically estimated average effect sizes for absolute and normalized knowledge gains were much smaller than the one found with posttest knowledge. However, and against our expectations, the correlations did not differ statistically significantly from zero. These results do not imply that the learners did not learn anything. The learners included in our meta-analysis acquired relatively large amounts of knowledge between pretest and posttest, as indicated by a Cohen's d of 1.62. However, averaged over all

included studies, individual differences in these knowledge gains were unrelated to individual differences in prior knowledge.

As expected, the correlations found with absolute knowledge gains are descriptively even lower than the ones found with normalized knowledge gains. Normalized gain scores acknowledge that learners with high prior knowledge have less room on improvement on the knowledge measure as compared to learner with low prior knowledge (Hake, 1998). Absolute gains scores neglect this fact and might lead to selective ceiling effects for learners with high prior knowledge. Put differently, absolute gain scores may underestimate their true knowledge gains, which is then also evident in a decreasing correlation between prior knowledge and absolute knowledge gains. Overall, the zero correlations found with absolute and normalized knowledge gains should not be interpreted, because there was a strong moderator effect leading to strong positive effect sizes in some studies and strong negative effect sizes in others, as we describe in the next section.

6.5.1.3 Moderator Effects

Our fourth research question concerned moderators of the correlations. We first discuss moderating effects on the correlation between prior knowledge and posttest knowledge and then moderating effects on the correlations of prior knowledge with absolute or normalized knowledge gains. The high number of included correlations r_P between prior knowledge and posttest knowledge allowed for many moderator analyses and gave a high statistical power to these analyses. We found that knowledge-related, learner-related, and environment-related variables as well as methodological study characteristics moderated the correlation. The knowledge-related characteristics were the correctness of knowledge, the knowledge type, the content domain, and the similarity of prior knowledge and posttest knowledge with respect to their content domain, modality, and knowledge type. Not surprisingly, correct knowledge was directly related and incorrect knowledge inversely related to positive learning outcomes. The finding that content domains differ in how strongly prior knowledge and posttest knowledge

correlate can be explained by assuming that content domains differ in the degree of their interrelatedness. Further research on how the conceptual structure of content domains affects the usefulness of prior knowledge is needed. The moderator effect of the similarity of prior knowledge and posttest knowledge converge with the widespread notion that the beneficial effects of domain-specific knowledge are domain specific, and that cross-domain transfer is difficult to achieve (Detterman & Sternberg, 1993; Hirschfeld & Gelman, 1994; Sala & Gobet, 2017). However, the correlations were still significantly greater zero when there was some dissimilarity, indicating that at least near transfer is difficult but possible to achieve. The only significant learner-related moderator was educational level, where the correlation significantly differed between primary education and higher education.

The only significant environment-related moderator was the cognitive demands of the instructional intervention. This moderating effect was so strong that it was found for posttest knowledge as the dependent variable as well as absolute and normalized knowledge gains, even though moderator analyses for knowledge gains had a low statistical power due to the small number of effect sizes. The cognitive demands of the intervention explained variance proportions of R^2 = .056 for posttest knowledge, R^2 = .281 for absolute knowledge gains, and R^2 = .366 for normalized knowledge gains.

Figure 6 visualizes the sets of correlations that were found in our meta-analysis for the overall effect, for interventions with high cognitive demands, and for interventions with low cognitive demands. We used example data of ten fictitious persons to visualize the correlation r_P^+ between prior knowledge and posttest knowledge, the correlation r_{NG}^+ between prior knowledge and normalized knowledge gains, and the Cohen's *d*-values of the knowledge changes from pretest to posttest. As shown in Figure 6, learners with higher prior knowledge also tend to have larger gains as compared to learners with lower prior knowledge in interventions with higher cognitive demands. Learners with higher prior knowledge tend to have smaller gains as compared to learners with lower prior knowledge in interventions with

lower cognitive demands. Together with the relatively high stability of the individual differences from pretest to posttest, this implies that interventions with high cognitive demands increase the variance of learners' amounts of knowledge and interventions with low cognitive demands decrease the variance.

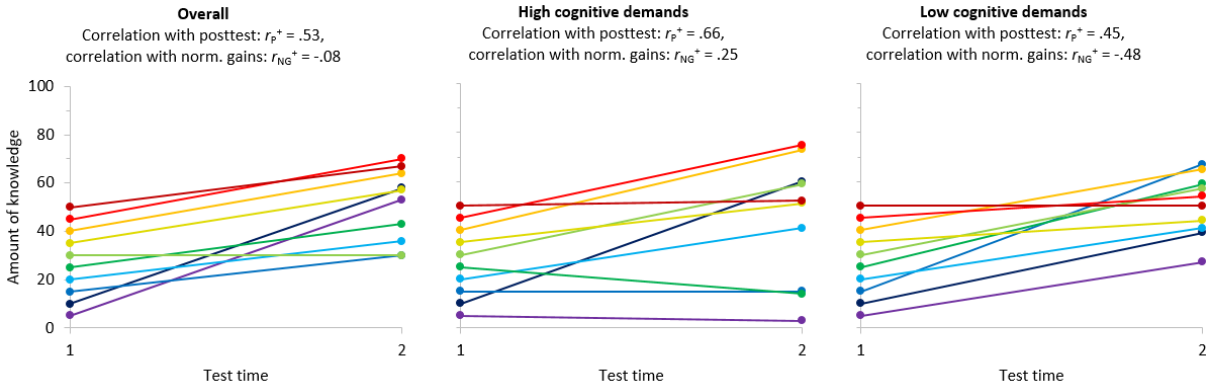

Figure 6. Correlations of prior knowledge with posttest knowledge and normalized knowledge gains found in the meta-analysis overall (left), for instruction with high cognitive demands (middle), and for instruction with low cognitive demands (right), visualized using ten fictitious persons' amount of knowledge before (T1) and after learning (T2) as example.

6.5.2 Limitations

This meta-analysis has at least four limitations. First, even though it included a wide and broad selection of studies, some content domains, age groups, and countries are under-represented. For example, few of the included studies were from the domains of chemistry, computer sciences, or medicine, few studies were from daycare or continued education, and we did not find studies from Africa and South America. This limits the generalizability of our results. Second, we focused on the quantity of prior knowledge only, but sometimes the quality of knowledge might be more predictive for future learning. Our meta-analysis might thus underestimate the size of the effect of prior knowledge on learning. Third, even though we scanned 5462 search hits for inclusion in our meta-analysis, the field of research on

knowledge acquisition is so broad, that there are surely many relevant studies not included in our meta-analysis. We see no reason to expect systematic differences between the studies included and not included. Likely, studies using the terms in our search string, and thus having been included in our meta-analysis, analyzed and reported the effects of prior knowledge more explicitly than other studies, what makes the included studies more useful for the meta-analytic aggregation and moderator analyses. Fourth, knowledge, as a multi-facetted and dynamically changing construct, can ideally be investigated by longitudinal multivariate methods, for example, latent transition analyses or latent growth curve analyses (Flaig et al., in press; Hickendorff et al., in press). However, the values of the path coefficients in multivariate statistical models depend on all other variables in the model and, thus, do not lend themselves for meta-analytic integration and interpretation. For these and other reasons, reporting standards for multivariate analyses recommend the publication of the zero-order correlations along with the multivariate findings (American Psychological Association, 2001, p. 161; Boomsma, 2000; Hoyle & Isherwood, 2013). We were able to include the bi-variate effect sizes from multivariate studies adhering to the reporting standards, but had to exclude the other multivariate studies from our analyses.

6.5.3 Implications

One central implication of our study for further research is the necessity to analyze both, the effect of prior knowledge on posttest knowledge, as well as the effect of prior knowledge on knowledge gains. These two groups of effect sizes are partly independent of each other, are of different magnitudes, and capture different aspects of the learning process. So far, studies have focused strongly on posttest knowledge. Future studies need to report the findings for posttest knowledge as well as knowledge gains, and compare them. Previous studies were also limited in that they mostly investigated correlations. Randomized controlled trials and longitudinal studies carefully controlling for confounding variables, such as

intelligence, are needed. Further, many studies measured prior knowledge and yet little is known about the actual causal effects of prior knowledge on learning.

A second implication of our study is that individual differences in knowledge are highly stable over the course of learning. In our meta-analysis, the correlation between prior knowledge and posttest knowledge was $r_P^+ = .525$, what is equivalent to a Cohen's *d* of 1.23. Hattie (2009) reported a meta-analytic rank-order of 138 variables associated with academic achievement, which did not include prior knowledge. Prior knowledge would be among the strongest three effect sizes of all 138 effects included in the rank-order. This demonstrates that domain-specific prior knowledge predicts individual differences in knowledge and achievement after learning better than almost all other variables. Assessments of prior knowledge, for example, in school entrance tests or in formative assessments can thus provide valuable information to teachers, parents, and the learners themselves.

The high stability of individual differences in domain-specific knowledge can be explained by assuming that the differences are aggregated over months or years of learning and are hard to overcome during short time periods, for example, in instructional interventions. This supports theoretical approaches emphasizing the long-term nature of domain-specific knowledge acquisition, for example, learning-trajectory approaches (Clements & Sarama, 2004), skill-building approaches (Bailey, Duncan, Watts, Clements, & Sarama, in press), theories of cognitive development (Siegler, 1996), and Ericsson's theory of expertise development (Ericsson & Charness, 1994).

Our results also shed light on the debate on Matthew effects in learning. A Matthew effect leads to larger learning gains for learners with high prior knowledge or achievement, thus widening the gap between low-prior knowledge learners and high-prior knowledge learners over time (Stanovich, 1986). Some previous studies found evidence in favor of Matthew effects in learning (Duff et al., 2015; Pfost, Dörfler, & Artelt, 2012), whereas others did not (Baumert et al., 2012; Schroeders et al., 2016). Our moderator analyses can explain this

heterogeneity (see Figure 6). Specifically, they show that a Matthew effect (a positive correlation between prior knowledge and learning gains) occurs only in instruction with high cognitive demands, for example, in which students have to reason, explain, and justify extensively. For instruction with low cognitive demands, where students have to listen, memorize, or practice simple tasks, prior knowledge and learning gains were negatively correlated, so that the learners with the lowest prior knowledge had the largest learning gains. This is sometimes also referred to as expertise reversal effect (Kalyuga, 2007) or compensation effect (Schroeders et al., 2016). Thus, teachers can widen or narrow achievement gaps between their students depending on how they design their instruction following their educational goals. The cognitive demands of the intervention moderated the effect sizes in our meta-analysis much stronger than the actual instructional methods (oral instruction, collaborative learning, instructional technology, etc.) used in the interventions. This demonstrates that the cognitive demands are not inherent to instructional methods, but that each method can be implemented in cognitively more or less demanding ways.

Our study exemplifies that knowledge sometimes is a more productive level of analysis than achievement in research on learning and instruction. Many studies in research on instructional effectiveness used achievement measures. However, achievement measures assess the learning outcomes of instruction of months or years and usually include several sub-domains, sub-skills, or competencies (OECD, 2016; Steinmayr et al., 2014). This breadth and complexity makes it hard to identify the sources of differences in achievement, for example, specific experiences and learning situations, concrete implementations of instructional methods, or mental processes in the learner (e.g. reasoning or elaboration). Compared to achievement, knowledge is a more homogeneous construct that can be changed within relatively short time frames by relatively simple interventions. This makes it much easier to identify the sources of knowledge changes than the sources of achievement changes. Because knowledge and achievement are closely related, it would be productive to trace

differences in achievement (e.g. between countries, schools, teachers, learners, or time points) back to the underlying differences in knowledge. These differences can then be traced back to the underlying classroom processes and mental processes of knowledge construction. In short, we cannot understand achievement without understanding knowledge.

Finally, our study demonstrates the need for the integration of the many findings on prior-knowledge effects on learning. Researchers from different sub-disciplines used varieties of research paradigms in hundreds of studies. This flood of information makes it hard to detect relations, imbalances, and open questions. In the following, we propose a framework designed to document central areas of research on prior knowledge and learning. The framework is deployable to summarize and integrate previous findings, to document empirical results and their interrelations, and to identify open questions for future research.

6.5.4 A Framework for Research on Prior Knowledge

The framework is based on the insight that the influence of prior knowledge on learning is a case of a moderated mediation (Preacher, Rucker, & Hayes, 2007), in which intra-learner processes mediate the effect of prior knowledge on cognitive learning outcomes, such as knowledge and achievement. These mediating processes are moderated by learner resources and the learner-environment interaction, which is in turn influenced by environmental resources.

Figure 7 shows the structure of the framework. The lists in each box provide examples and are not exhaustive. The framework is based on the following five assumptions:

1. In research on prior knowledge, it is useful to distinguish between knowledge (bottom level in Figure 7), the learner (middle), and the environment (top). It is also useful to distinguish between dynamic processes affecting knowledge construction (right), and more stable resources moderating these processes (left). By stable, we mean that the resources usually change at a much slower rate than knowledge. Thus, they are more likely to affect knowledge construction than they are to be affected by knowledge

construction. This is indicated by the one-headed arrows leading away from the resources, for example, from learner resources to intra-learner processes.

2. Of all framework components, only intra-learner processes directly mediate the relation between prior knowledge and learning outcomes, because only the learner her- or himself can recall prior knowledge from long-term memory and can encode new knowledge in long-term memory.
3. Relatively stable learner resources moderate the relation between prior knowledge and learning only to the extent with which they affect the processing of knowledge in the learner.
4. Relatively stable environmental resources, for example, pre-prepared learning environments such as school instruction, can affect the relation between prior knowledge and learning outcomes only to the extent with which the learner interacts with these environments and to the extent with which these interactions affect the processing of knowledge in the learner. Therefore, learning environments can only provide learning opportunities but cannot guarantee what and how much is learned. Instead, the learning outcomes depend on how the learner adopts these opportunities (e.g., Schrader & Helmke, 2015).
5. Knowledge is a component of achievement (OECD, 2016; Steinmayr et al., 2014) and related cognitive-learning outcomes (Ericsson & Charness, 1994). Therefore, prior knowledge and the other parts of the framework do not only affect the acquisition of knowledge but also the acquisition of cognitive learning outcomes in general, including, but not limited to, achievement, competence, and expertise.

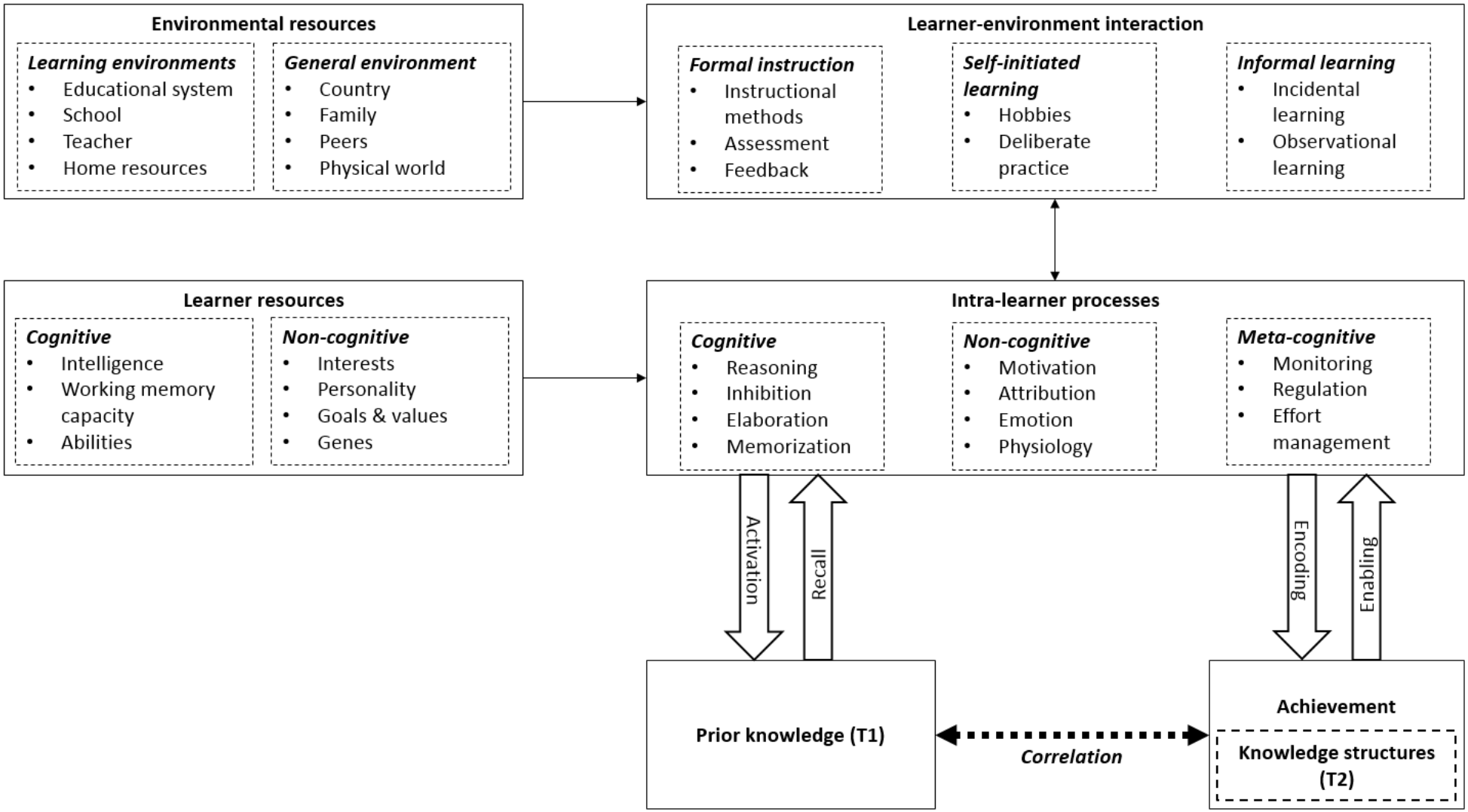

Figure 7. Framework for research on the relation between prior knowledge and learning. The effect of prior knowledge on knowledge and achievement is mediated by intra-learner processes, that are moderated by learner resources and the learner-environment interaction.

This framework describes how the use of prior knowledge is embedded in the wider context of learning and instruction. Whereas other frameworks have been published that partially included similar assumptions (Biggs, 2003, p. 19; Schrader & Helmke, 2015; Ullén et al., 2016), the innovative and defining key characteristic of our model is that it describes the effect of prior knowledge on learning as moderated mediation and distinguishes between mediating variables, direct moderators, and indirect moderators of the mediation. The framework thus links cognitive and educational research on knowledge construction to research on moderated mediation, which is an active field of ongoing methodological research (Preacher et al., 2007). The relational structure of the framework also demonstrates the importance of future randomized controlled trials and careful process analyses to investigate the processes mediating the effect of prior knowledge on learning and how these are moderated by knowledge-, learner-, and environmental characteristics.

6.5.5 Conclusion

Many studies investigated the role of prior knowledge in learning, but most of them only reported how prior knowledge, used as a covariate, correlates with learning outcomes. Our meta-analysis demonstrates that this correlation is positive and strong – in general, as well as for many levels of knowledge-, learner-, and environment related moderators. Thus, individual differences in knowledge are highly stable, so that prior knowledge, measured before learning, is of a high predictive power for knowledge and achievement after learning. Randomized controlled trials indicate that this relation is causal rather than merely correlational. Only few studies investigated the association of prior knowledge with knowledge gains. Overall, this correlation was zero. However, it was positive for interventions with high cognitive demands and negative for interventions with low cognitive demands. Thus, we found a Matthew effect (i.e., a widening gap between learners' amounts of knowledge over time) for instruction with high cognitive demands and a compensation effect or expertise reversal effect (i.e., a closing gap between learners' amounts of knowledge

over time) for instruction with low cognitive demands. Future research might benefit from using more randomized controlled trials, analyzing learning gains as well as posttest performance, and a systematic evaluation of mechanisms that mediate and moderate the influence of prior knowledge on learning. Ausubel (1968, p. vi) broadly hypothesized that the "important single factor influencing learning is what the learner knows already" – our meta-analytic results allow refining this hypothesis scaffolded by our framework for further scrutiny of the effects of prior knowledge.

7. Study 3: Conceptual Change and Knowledge Integration as Learning Processes in Higher Education: A Latent Transition Analysis[1]

7.1 Abstract

Conceptual change, that is, a restructuring of incompatible prior knowledge, is a well investigated learning mechanism in school children's acquisition of new concepts. An understanding of academic concepts is also a central learning goal of higher education. However, there is almost no research on conceptual change and knowledge integration in higher education. We tracked 137 undergraduate psychology students' concepts of human memory longitudinally over four semesters. A latent profile transition analysis (LPTA) showed that the students' development followed six transition paths between four knowledge profiles. These developmental pathways were well-ordered, indicated a general trend from fragmented knowledge to integrated scientific knowledge, and correlated with the students' university grades and with an additional test of memory understanding. The findings highlight the importance of conceptual change, in particular, knowledge integration, in higher education and exemplify the usefulness of LPTA for modeling individual differences in longitudinal changes of multidimensional knowledge structures.

7.2 Theoretical Background

7.2.1 Introduction

Students' understanding of academic concepts, for example, force in Physics, supply and demand in Economy, or human memory in Psychology, is a central learning goal of higher education. Conceptual understanding is a cornerstone of professional expertise (Tynjälä, 1999). It helps learners make predictions, explain observations, reason about the interrelations of facts, infer new knowledge, choose between alternative procedures, and construct new

[1] The author of this dissertation is the second author of the presented study.

problem solving strategies (Goldstone & Kersten, 2003; Machery, 2010; Rittle-Johnson et al., 2015). Accordingly, the European Qualification Framework for Lifelong Learning lists "advanced knowledge of a field of work or study, involving a critical understanding of theories and principles" (European Commission, 2008, p. 12) as a central qualification for reaching a Bachelor's degree.

In many cases, conceptual change, that is, a restructuring of the learner's prior knowledge, is a necessary part of acquiring new conceptual knowledge (Carey, 1985; Posner, Strike, Hewson, & Gertzog, 1982; Vosniadou, 2008). Prior knowledge guides and constraints the interpretation of new knowledge and its encoding in memory. It often stems from observations and explanation attempts in everyday life outside formal instruction and thus can be incompatible with the scientific concepts to be learned (Carey, 1992). This explains why acquiring an understanding of academic concepts can be so difficult. Conceptual change is investigated by educational, developmental, cognitive, and philosophical scientists, who found the approach productive in content domains as diverse as physics, chemistry, biology, mathematics, medicine, and the social sciences (Michael Schneider, Vamvakoussi, & Van Dooren, 2012; Vosniadou, 2008). Some of the past findings have direct implications for the design of effective learning environments, for example, school instruction (Duit, Treagust, & Widodo, 2008), professional development programs for teachers (e.g., Hewson, Tabachnick, Zeichner, & Lemberger, 1999), and projects to foster instructional quality in schools (Beeth et al., 2003).

Empirical research on conceptual change of students in higher education is virtually non-existent, despite the importance of academic concepts as learning goals in higher education. Almost all studies on conceptual change focused on students in K-12 schools or even younger children. In the present study, we examined to what extent conceptual change is still a relevant learning mechanism in higher education and whether it leads to similar developmental patterns in higher education as it does in K-12 school learning. We expected

conceptual change to still be relevant in higher education, because learning by conceptual change has been described as a general human learning mechanism that is relevant independently of age group and content domain. For example, there are some similarities between children's thinking processes when acquiring new concepts and scientists' thinking processes when developing a new theory, which hints at an age-general importance of conceptual change (Gentner et al., 1997).

7.2.2 Knowledge Fragmentation and Integration

Research with school-aged children found that the fragmentation and integration of knowledge are two central component processes of conceptual change (Michael Schneider & Hardy, 2013). Networks of conceptual knowledge in long-term memory can comprise types of elements, such as observations, hypotheses, explanations, analog mental models, mental images, category exemplars, and subjective theories (Goldstone & Kersten, 2003; Machery, 2010). These elements have been acquired in situation as diverse as conversations with peers, everyday-life observations, internet videos, school instruction, books, or movies. Learners do not always understand how these different and sometimes even conflicting pieces of knowledge relate to each other, and store them independently in long-term memory. This leads to fragmented knowledge.

Another source of knowledge fragmentation is the fact that storing correct concepts in long-term memory does not necessarily erase related misconceptions. Converging evidence from reaction times studies with sentence verification tasks (Potvin, Masson, Lafortune, & Cyr, 2015; Shtulman & Valcarcel, 2012), multiple choice tests (Hardy, Jonen, Möller, & Stern, 2006), and interviews (diSessa et al., 2004) shows that naïve misconceptions and scientifically correct concepts or parts thereof can co-exist in long-term memory and do so frequently, not only in children, but also over the life-span (Shtulman & Harrington, 2016). Pieces of fragmented knowledge are activated dependent on the context (diSessa et al., 2004), so that learners do not necessarily realize when they hold pieces of knowledge in long-term

memory that support or contradict each other.

Thus, the integration of pieces of knowledge into a coherent overarching knowledge structure is an important aim of instruction (Linn, 2006; Michael Schneider, 2012). This can include connecting previously isolated pieces of knowledge in memory and subsuming previously unrelated concepts under a general principle. Understanding these relations can help students to differentiate better between normatively correct and incorrect ideas, thus, leading to more homogeneous and more correct knowledge. Teachers can stimulate knowledge integration by eliciting students' knowledge, adding new normative concepts, helping students to develop criteria for evaluating alternative conceptions, and by encouraging students to compare the alternatives and to sort out inadequate conceptions (Linn, 2006).

7.2.3 A Latent Profile Transition Analysis of Fragmented and Integrated Knowledge

Developmental patterns of co-existing pieces of knowledge can be investigated by latent profile transition analyses (LPTA), as demonstrated by three studies with school children on knowledge development in mathematics and science (Hardy et al., 2006; Kainulainen, McMullen, & Lehtinen, 2017; McMullen, Laakkonen, Hannula-Sormunen, & Lehtinen, 2015). Michael Schneider and Hardy (2013) investigated third-graders' understanding of floating and sinking of objects in liquids. The children participated in several sessions of either (1) a constructivist learning environment with a high degree of instructional support given by the teacher, or (2) a constructivist learning environment with a low degree of instructional support, or (3) a baseline control group without instruction on the topic (see Hardy et al., 2006, for details of the interventions). Before and after the instruction as well as one year later the children indicated in a multiple choice test how strongly they agreed with a number of statements. Each statement described (a) a common misconception, (b) an everyday life explanation, which has some explanatory power in everyday life but does not hold up to systematic scientific evaluation, or (c) the relevant scientifically correct concepts.

The three sum scores indicating how often each child agreed with misconceptions, everyday conceptions, or scientific concepts were used as indicators of latent profile memberships at the three measurement points in a latent profile transition model. The parameters of this model were estimated so that the similarity of the scores of persons grouped in the same latent profiles was maximized and the similarity of the scores of persons grouped in the different latent profiles was minimized (see Hickendorff et al., in press, for methodological details). Thus, the latent profiles indicated groups of learners who had the same configuration of misconceptions, everyday conceptions, and scientific concepts. In addition to the profile characteristics and memberships at each measurement point, the model estimation also yielded the frequencies of the profile transitions over time. These were interpreted as pathways of conceptual change.

The model results in the study by Michael Schneider and Hardy (2013) indicated five latent profiles. Some of these had mean score profiles that indicated integrated knowledge, that is, high profile mean scores for misconceptions only or for scientific concepts only. Other profiles indicated fragmented knowledge. In these profiles two or three of the scores for misconceptions, everyday concepts, or scientific concepts were higher than the sample mean. Most participants were on one of seven developmental pathways between these five profiles over time. These transition paths indicated a general trend from misconceptions and fragmented knowledge towards more correct and integrated knowledge. However, there were strong individual differences. Knowledge fragmentation increased on some paths and decreased on others. About 20% of the children still had fragmented knowledge, that is, co-existing misconceptions, everyday concepts, and scientific concepts, even one year after participating in the constructivist learning environment. The instructional condition was related to the frequency of the transition paths. The constructivist learning environment with a high degree of instructional support led to the most integrated knowledge and the untreated control group to the least.

The other two studies using latent transition analyses to investigate knowledge development traced school students' knowledge of rational numbers over time (Kainulainen et al., 2017; McMullen et al., 2015). Similar to Schneider and Hardy's findings, students' knowledge of rational numbers was captured by a small number of knowledge profiles and systematic transition paths between these profiles over time, some of which could be interpreted in terms of conceptual change. However, the generalizability of these findings to other age groups and content domains remains unclear (cf. Edelsbrunner, Schalk, Schumacher, & Stern, under review; McMullen, Van Hoof, Degrande, Verschaffel, & Van Dooren, under review)

7.2.4 Is Conceptual Change Still Relevant in Higher Education?

In the current study, we used latent profile transition analysis to investigate whether conceptual change and, more specifically, knowledge fragmentation and integration still can be found in higher education and whether their quality is related to the learning outcomes, that is, students' grades. There are at least three reasons to expect that this might not be the case. First, students in higher education successfully participated in school instruction, perhaps leading to a firm fundament of correct and integrated knowledge that higher education can build on. Second, students in higher education tend to have better developed meta-cognitive strategies than school children (Weil et al., 2013). Thus, they might be able to monitor, identify, and understand the confirmatory or contradictory relations between their prior knowledge and the knowledge to be learned making knowledge integration a trivial process. Finally, arguably, the content of higher education programs is more abstract than the content of school lessons, so that less interference between prior knowledge from everyday life and the new knowledge may occur.

On the other side, there are also a number of reasons why conceptual change could still affect student learning in higher education. First, the structure and the functions of human memory do not change fundamentally between K-12 school and higher education. Adults'

conceptual knowledge is still organized as a network, so restructuring this network might be necessary when prior knowledge and new knowledge are incompatible. Second, empirical research shows that university students still have many misconceptions, which are stable, hamper subsequent learning, and require restructuring (Merz, Dietsch, & Schneider, 2016; Shtulman & Valcarcel, 2012). Third, a recent meta-analysis on undergraduate science course innovations found that so-called conceptually-oriented learning tasks had a substantial effect on student achievement. Averaged over nine studies the effect size was $d = 0.47$ ($SD = 0.70$). The authors defined that tasks are conceptually oriented when they "elicit students' level of understanding of key science concepts, […] engage students in conceptual schemes within a topic rather than isolated facts, […] and engage students with real-world problems in creative ways that reflect a conceptually integrated understanding of the content" (Ruiz-Primo, Briggs, Iverson, Talbot, & Shepard, 2011, p. 1269). Based on these findings we hypothesize that our latent profile transition analysis will find evidence for conceptual change and, in particular, for knowledge fragmentation and integration, in higher education.

7.2.5 Psychology Students' Concepts of Human Memory

We tested our hypotheses in a longitudinal study on Psychology students' understanding of human memory. Specifically, we assessed at four measurement points in the first four semesters of a Bachelor program in Psychology to what extent the students had the misconception of memory as a place for the static storage of information and to what extent they had the correct concept of memory as a dynamic system involving the construction and re-construction of knowledge (Lynn & McConkey, 1998). Examples of memory processes that modify the information to be stored are interference (Bjork, 1992), chunking (Gobet et al., 2001), and source monitoring (M. K. Johnson, Hashtroudi, & Lindsay, 1993).

Human memory is a complex causal system and one of the central constructs of Psychology (Baddeley et al., 2009). The search term *memory* has more than 190.000 hits in the literature database PsycInfo (November 2016). These articles are from fields as diverse as

cognitive, educational, developmental, clinical, social, forensic, and biological psychology. An understanding of memory properties and its processes is also important in numerous psychological professions, and a flawed understanding of human memory in professionals who work in the clinical or legal context can have negative consequences (Garry, Loftus, & Brown, 1994). As far as we know, there is no standardized test of students' concepts of human memory. Therefore, we used a self-developed test in our study.

To assess the criterion validity of our new test, we additionally presented the participants with the implicit memory theory scale (IMTS; Niedźwieńska, Neckar, & Baran, 2007), which tests how skeptic individuals are with respect to the credibility of autobiographical memory. We expect that students who have a concept of human memory as a dynamic system that constructs and re-constructs information will be more skeptical with respect to the validity of autobiographic memories than other students.

7.2.6 The Current Study

In sum, conceptual understanding is a central learning goal of higher education. Yet there is almost no published research on conceptual change in higher education students. For this reason, we used a latent profile transition model with longitudinal data from university students in the current study. We constructed the measures so that the latent profiles can be interpreted in terms of knowledge fragmentation and integration and that any profile transition can be interpreted as signs of conceptual change. We tried to replicate Michael Schneider and Hardy (2013) findings on a general conceptual level, in particular, the trend from misconceptions to scientific concepts, the trend from fragmented to integrated knowledge, and the persistence of fragmented knowledge in some learners even after instruction.

Based on the previous findings of studies with school children we had the following hypotheses for our study in higher education: (1) Persons differ in their knowledge profiles and transition paths. However, the number of profiles and paths is relatively small. (2a) There are transitions between latent profiles differing in their knowledge profiles over time, giving

evidence of conceptual change in higher education. (2b) The transitions reflect an overall developmental trend from profiles with higher scores of misconceptions towards profiles with higher scores of scientific concepts. (3a) At least one of the knowledge profiles indicates fragmented knowledge by indicating high agreement with mutually incompatible concepts. (3b) The profiles indicating fragmented knowledge will become less frequent but will not disappear completely over time. (4) The latent profiles differ in their mean scores on the Implicit Memory Theory Scale indicating that students' profiles of knowledge of human memory are related to how much trust they put in autobiographic memory. (5) The latent profiles differ in their grades on the human memory course showing that students' profiles of knowledge of human memory are related to grades in higher education.

7.3 Method

7.3.1 Participants

137 students enrolled in a Bachelor of Psychology program at a mid-sized university in a mid-sized German city participated at T1 in our study. Of these, N = 126 participated again at T2, N = 116 at T3 and N = 115 at T4. Almost all participants were German, and all were fluent speakers of German. At T1, the sample mean age was 20.4 (*min* = 18; *max* = 31) and all students were at the beginning of their first semester in the program. About 82% of participants in the sample were female. This proportion is slightly higher than the proportion of all females in the program, which was 60-70%. Participation in the study was anonymous and voluntary. To keep dropout at a minimum, participants were financially compensated with €25 per wave of the longitudinal study and an additional completion bonus of €100 if they had participated in all waves.

The study was conducted in full accordance with the Declaration of Helsinki and the APA Ethics Code (American Psychological Association, 2002). Prior to their participation, students received an informed consent form including, among others, the following

information (based on the APA Ethics code): (1) a statement on the purpose of the research as well as the expected study duration and procedures, (2) a statement that participation is voluntarily and that it may be terminated at any point; (3) a statement that there are no potential risks, discomfort or adverse effects with regard to their participation; (4) a statement that data is collected anonymously, and that even though some personal data (e.g., e-mail addresses) will be collected for organizational purposes, these data will not be used to identify individual participants and will be deleted as soon as possible.

7.3.2 Procedure

The students were tested longitudinally at four measurement points covering the first four semesters of the Bachelor in Psychology program. Data collection took place between November 2013 and May 2015. Baseline data collection (T1) took place during the first six weeks of the participants' first semester in the program, followed by three consecutive waves of measurement (T2, T3, T4) at the beginning of the second, third, and fourth semesters, respectively. Each wave consisted of two parts: A home module and a subsequent lab module. The home modules included several self-report measures and were completed online under unsupervised conditions before the respective lab sessions took place; answering the questionnaires took between 30 and 50 minutes (dependent on the wave). In the lab modules, knowledge and achievement tests were conducted in the university's computer labs. Groups of 1 to 25 participants were supervised by student experimenters and completed the tests individually and at their own pace. Each lab module took about 120 minutes.

7.3.3 Measures

7.3.3.1 Concepts of Human Memory

Conceptual understanding of human memory was assessed by a self-constructed multiple choice test at the beginning of each of the four lab modules. Previous studies (cf. Michael Schneider & Hardy, 2013) have found that students' answers to multiple choice items are well

in line with their answers to interview questions that aim to assess the same concepts. The tasks were presented at a computer screen and the answers were entered by mouse clicks. Each of the nine test tasks began with a written description of a situation in which a certain memory process is particularly important. These situations were derived from and similar to classical experiments in memory research (Brooks, 1967; Chase & Simon, 1973; Ericsson et al., 1980; Goff & Roediger, 1998; Henik & Tzelgov, 1982; Hovland & Weiss, 1951; Loftus, 1975; Melcher & Schooler, 1996; W. Schneider et al., 1989). In three situations, interference between elements in memory (Bjork, 1992) can be expected, in three tasks the chunking of elements into bundles of information (Gobet et al., 2001; M. K. Johnson et al., 1993), and in three tasks source monitoring mechanisms(M. K. Johnson et al., 1993). Each of the nine publications was cited between 181 and 3286 times, had been replicated repeatedly, and was (still) widely accepted by the scientific community.

In each of the nine tasks, the learners were presented with six statements about memory processes that might be relevant in this situation (see the example task in Figure 8 and Appendix H for examples of all three types of tasks). There were three types of statements: (a) *Correct processing statements* described the memory process actually occurring in that situation as firmly established by empirical research, as described in the previous paragraph. These statements were compatible with the view of memory as a dynamic system involving the construction of knowledge. (b) *Incorrect processing statements* also describing dynamic knowledge construction processes in memory, however, only processes not occurring or being irrelevant in that specific situation as shown by established empirical research. Incorrect processing statements were partly correct (because they describe memory as a dynamic system) and partly incorrect (because they described processes not relevant in the respective situation). (c) *Static storage statements* described memory in that situation as a place for the static storage of information, that leaves the nature and the content of the stored information unchanged. They do not conform to the established empirical findings and the view of the

scientific community and can be seen as misconception.

Wild horses

Two groups of children participate in the study.

The children in one of these groups (Group A) are twelve year-olds, who know little about horses in general.

The children in the other group (Group B) are eight year-olds, who know a lot about horses in general.

The two groups do not differ in terms of the children's intelligence and the number of boys and girls in the group.

All participants read a simple text about wild horses that teaches them what kinds of species there are, where they are found and what they need to live. Afterwards the children are asked to answer question assessing their knowledge and comprehension based on the text from memory.

How sure are you for each of the following statements that it is correct or incorrect?

	Definitely incorrect	Definitely correct
Group A will perform markedly better than Group B because they have four more years of practice in reading and remembering from texts than the younger children.	O O O O O O O	
Group B will perform markedly worse than Group A, because the children are on a lower stage of cognitive development and therefore cannot process information so well.	O O O O O O O	
Group B will perform markedly better than Group A because they have more prior knowledge and therefore can store information from the text in a more structured way.	O O O O O O O	
Group A will perform markedly worse than Group B because the memory of twelve year-olds is partly impaired already due hormonal changes during puberty.	O O O O O O O	
Both groups will perform almost equally well because they have read the same text with the result that each person has memorized the same information.	O O O O O O O	
There barely will be any difference between the two groups because their intelligence is similar and therefore they can process information almost equally well.	O O O O O O O	

Figure 8. Example task in the test of concepts of human memory (in this figure translated from German to English). The correct answer is in bold print.

The participants evaluated each statement on a seven-point rating scale from definitely incorrect to definitely correct. In each task, only one memory process was relevant. Therefore, there was one correct processing statement, but three incorrect processing statements, and two static storage statements for each task. Half of the statements used a negative wording, e.g., "The two groups do not differ". The order of the nine tasks and of the six statements in each task was randomized separately for each person. We reversed the scores for the incorrect statements and then computed a separate sum score for static correct processing statements, incorrect processing statements, and static storage statements, respectively.

The test was not tailored to the content of the specific course. Human memory is referred

to in many courses of the Bachelor Psychology program. So we developed the test to reflect students' accumulating knowledge of memory over several semesters and courses. Still the test content is more similar to the lecture "Human Memory" than to any other course in the students' program. Students typically attend this lecture during their first semester in the program.

7.3.3.2 Implicit Memory Theory Scale

At T3, we administered a German translation of the Implicit Memory Theory Scale (IMTS; Niedźwieńska et al., 2007) as part of the at-home module. One researcher from our lab translated it from English into German. To check for the validity of the translation, a second researcher from our lab translated the German items back into English without knowing the original English items. In case of disagreements between the two English versions, the German items were modified by both researchers together. The IMTS is a standardized instrument to test how individuals judge the credibility of autobiographical memory. Low skepticism is associated with misconceptions of human memory, such as the belief that memory is a static storage and that recollection is an accurate representation of real events, whereas high skepticism is associated with the belief that memories are reconstructed during recall and that memories are prone to qualitative changes. The global original scale has good psychometric properties (Niedźwieńska et al., 2007). In the original study, the internal consistency was good with a Cronbach's alpha of .83 and the retest reliability was also good, with a value of .81 for a two-week interval and .74 for a seven-month interval. The IMTS proved to be sensitive to knowledge differences between psychologists and non-psychologists, as well as to knowledge differences between psychology students before and after instruction on autobiographical memory. We report the mean and standard deviation of the IMTS scores in our sample in Table 6.

7.3.3.3 Grades

At T4, the university administration sent us the grades for all the exams the participants in our study had completed so far. In the German grading system, grades range from 1 to 5, with smaller numbers indicating better performance. The students are relatively free in choosing the sequence of their exams, so the types and numbers of completed exams varied between participants. For our analysis, we only used the grades of a course in general psychology, which covers the topics *human memory*, *learning*, *motivation* and *emotion*, and thus, is the best proxy for human memory competence. At T4 82 % of the sample had completed the respective exam. The mean and standard deviation of the grades is reported in Table 6.

7.3.4 Statistical Analysis

In the latent profile transition model, we specified a latent profile variable for each of the four measurement points. Each latent profile variable had the sum scores for correct processing, incorrect processing, and static storage at the respective measurement point as indicators of the persons' profile memberships. The profile means of the indicator variables were constrained to be equal over time, so that fewer parameters had to be estimated and that the profiles had the same interpretation at all four measurement points. The number of participants in each profile and the variance of the indicator variables were allowed to differ between measurement points. The profile membership at each time point was used as a predictor of the profile membership at the respective next time point, so that intercepts and regression weights of three multinomial logistic regressions were estimated. MPlus uses these parameters to compute transition probabilities and transition paths between the latent profiles over time in the model. Overall, the model has 63 free parameters. These are the 12 means and 12 variances of the class indicators (i.e., for the three indicators at the four time points), 12 regression intercepts (i.e., for four minus one latent classes at four points in time), and 27 regression weights (for regressions of four minus one profiles at one time point to four minus

one profiles at the respective next time point times three pairs of time points).

The data were prepared for structural equation modelling using SPSS IBM® Version 23. The latent profile transition model was analyzed in MPlus 7.31 (Muthén & Muthén, 1998-2015) We used the maximum likelihood estimator with robust standard errors (MLR), which is the default estimator for latent transition models in MPlus. The model was estimated using the MPlus default handling of missing data according to which all available data is used to estimate the model (Muthén & Muthén, 1998-2015). When there was missing data, the covariance coverage of the variables used in our analysis was between .825 and .920 which far exceeded the minimum covariance coverage of .100. In latent profile, latent class, and latent transition analysis it is important to make sure that the best log-likelihood value is replicated several times and to avoid improper solutions due to local maxima. We therefore increased the default number of random starts in the initial stage to 400 sets and the number of final stage optimizations to 100. We used unstandardized (raw) scores of the indicator variables for the model estimation to yield unbiased model results. These scores are reported in Table 6. After model estimation and to aid the interpretation of the results, we standardized the profile mean values and standard deviations to T-scores, which have a mean of 50 and a standard deviation of 10. This was done by first subtracting the sample mean from the profile mean and dividing the result by the sample standard deviation separately for each profile indicator in each profile. The sample means and SDs were M = 0.86, SD = .48, for static storage, M = 1.66, SD = .59, for incorrect processing, and M = 4.22, SD = .65, for correct processing. Following, the resulting values were multiplied by 10 and added to 50. We report these standardized scores in the following to aid interpretation.

Table 6

Descriptive statistics of all unstandardized scores

	M	*SD*	*N*
Static Storage T1	1.19	.609	137
Incorrect Processing T1	2.00	.599	137
Correct Processing T1	3.79	.715	137
Static Storage T2	.798	.591	126
Incorrect Processing T2	1.66	.654	126
Correct Processing T2	4.41	.738	126
Static Storage T3	.736	.628	116
Incorrect Processing T3	1.49	.693	116
Correct Processing T3	4.41	.823	116
Static Storage T4	.584	.579	114
Incorrect Processing T4	1.32	.714	114
Correct Processing T4	4.49	.859	114
IMTS T3	31.0	10.2	116
Grades T4	1.99	.897	111

7.4 Results

In the following three subsections, we first describe how we determined the number of latent profiles, then characterize the latent profiles and describe the transition paths between the latent profiles over time. Finally, we analyze the relationship between the latent profiles and outside criteria, that is, the Implicit Memory Theory Scale and course grades.

7.4.1 Determining the Number of Latent Profiles

We selected the number of latent profiles based the recommendations given by Nylund, Asparouhov, and Muthén (2007). The best fitting model is determined by repeatedly estimating the model with increasing numbers of classes or profiles and comparing their model fits. There is no definite standard for deciding on the number of latent profiles (Nylund et al., 2007). Better-fitting mixture models are characterized by lower comparative fit indices, i.e. lower values in the Akaike information criterion (AIC), Bayesian information criterion

(BIC), and sample-size adjusted BIC. Another way to examine model fit is to compare the classification quality of individuals into latent profiles between models that differ in the number of profiles assumed. Classification quality or *entropy* is measured on a scale that ranges between zero and one, with a value of one indicating perfect classification of individuals into latent profiles (Clark & Muthén, 2009).

Table 7

Fit indices for the latent transitions models with one to five latent profiles

Index	One profile	Two profiles	Three profiles	Four profiles	Five profiles
Log Likelihood	-1513	-1350	-1278	-1245	-1212
Free parameters	24	25	41	63	91
AIC	3074	2749	2638	2617	2607
BIC	3145	2822	2758	2801	2873
Sample-size adjusted BIC	3069	2743	2628	2601	2585
Entropy	-	0.84	0.83	0.86	0.84

Note. AIC = Akaike information criterion; BIC = Bayesian information criterion.

We repeatedly estimated the model, each time with a different number of specified latent profiles. For each model, the best log-likelihood value was replicated several times with the initial specification of 400 sets of random starts and 100 iterations, that is, 99 times in the one-class model, 100 times in the two-class and three-class model, 25 times in the four-class model, and 8 times in the five-class model. In the second step, in which we ran the models again with twice the number of random starts and final stage iterations, each of these values was replicated again. Table 7 presents the model fit indices for the models with one to five latent profiles. The models with two to five latent profiles had entropies above .80, which is high according to Clark and Muthén (2009). The BIC was lowest for the three-profile model,

whereas the AIC and sample size-adjusted BIC were lowest for the five-profile model, and entropy was highest in the four-profile model. Therefore, all three models fit the data approximately equally well. We chose the four-profile model because the four-profile model but not the three-profile model included a profile indicating strong misconceptions (high values for static processing only). Even though only small sample proportions showed this profile in our study, it is still interesting, because it corresponds to the lowest level of knowledge of human memory, thus is the starting point of competence development in a domain, and had also been found in previous research (cf. Michael Schneider & Hardy, 2013). Furthermore, we chose the four-profile model over the five-profile model because it was more parsimonious and had a lower BIC. All subsequent analyses were conducted with the four-profile model.

7.4.2 Interpretation of the Profile Means

Table 8 lists the *T*-standardized indicator means for the four latent profiles. In Figure 9 the standardized scores are represented graphically. To aid interpretation, we tested for static storage statements, incorrect processing statements, and correct processing statements in each latent profile at each measurement point whether the latent profile mean significantly deviated from the overall sample mean. For example, we fixed the mean of static storage in profile 1 at the overall sample mean of static storage while fixing the other means at the values of the initial 4 x 4 latent transition model. In total, we computed 3 x 4 = 12 models for the three mean scores across the four latent profiles. Then, we computed log likelihood ratio chi-square difference tests to test for the difference between the fixed mean models and the initial model. We computed the robust test statistic for MLR estimation (Satorra & Bentler, 2010) as reported by Asparouhov and Muthén (2010; formula 2) and adjusted the significance level for the number of comparisons using the Bonferroni method, yielding a significance level of $\alpha < \frac{.05}{12} < .004$. Table 8 shows the results of these comparisons.

We interpreted the latent profiles based on the standardized mean scores, the results of the computed log likelihood ratio chi-square difference tests, and assigned labels to them (see column 2 in Table 8). We labeled the profile C1 misconceptions profile because members of this profile show the highest scores for the static storage misconceptions of all classes, average scores for incorrect processing statements and below-average scores for correct processing statements, even though the first and last deviation were not significant in the chi square difference test. However, it should be noted that only a small number of individuals were assigned to the misconceptions profile, yielding large standard errors and the Bonferroni correction is a rather conservative correction of the significance level. We called the profile C2, which shows significant above-average means on two of the three scales, fragmented profile because these learners express inconsistent knowledge: On the one hand, they believed that human memory was a static storage and on the other hand they believed that memory processes and re-constructs information.

This profile supports our hypothesis H3a, that some learners have fragmented knowledge rather than integrated knowledge. The third profile, C3, indicated that the participants mostly chose the answer category in the middle of the rating scale with average scores on all three scales. We labeled it indecisive profile. Finally, we labeled the profile of C4 scientific profile because it displayed significant above-average means of scientific concepts and significant below-average scores of static storage statements and incorrect processing statements, representing the ideal learning outcome.

Table 8

Assigned labels for the latent knowledge profiles, sample proportions, standardized scores of the scales' means, and significance of the deviation of each mean from 50

	Label	Sample Proportion in %				Static storage		Incorrect Processing		Correct Processing	
		T 1	T 2	T 3	T 4	*M*	*p*	*M*	*p*	*M*	*p*
C1	Misconceptions profile	7	5	6	4	92	.053	49	1.00	30	.544
C2	Fragmented profile	56	22	17	17	71	<.001*	71	<.001*	39	.028
C3	Indecisive profile	37	46	43	38	40	.281	48	.927	52	.998
C4	Scientific profile	0	27	33	42	26	<.001*	25	<.001*	70	<.001*
	Total	100	100	100	100						

Note. * significant on the level of $\alpha < 0.004$, corrected for multiple testing.

To further aid interpretation, we tested whether the latent profiles differed significantly in their means on the three indicator variables. We computed a separate model for each pair of the four latent profiles. In each model, we constrained two class profiles to be equal whereas all other model parameters were fixed to the values found with the original (unconstrained) four-profile model. We corrected for multiple testing using the Bonferroni method and set the significance level at $\alpha < \frac{.05}{6} < .008$. Five of the six constrained models had a significantly worse fit than the original four-profile model, as found by likelihood ratio chi-square difference tests, all *p*s $\leq .001$. Only the comparison between the unconstrained model and the model in which the first and the third profile were constrained equal was not statistically significant, $p = .012$. The reasons for this comparison missing the critical significance level was the low frequency of the misconceptions profile and the very conservative Bonferroni-corrected critical significance level. Descriptively, as shown in Figure 9, there is a large difference between these two profiles. Thus, the results indicate that all four class profiles are mutually different and support our first hypothesis, that there are latent profiles of learners differing systematically in their conceptual knowledge (H1).

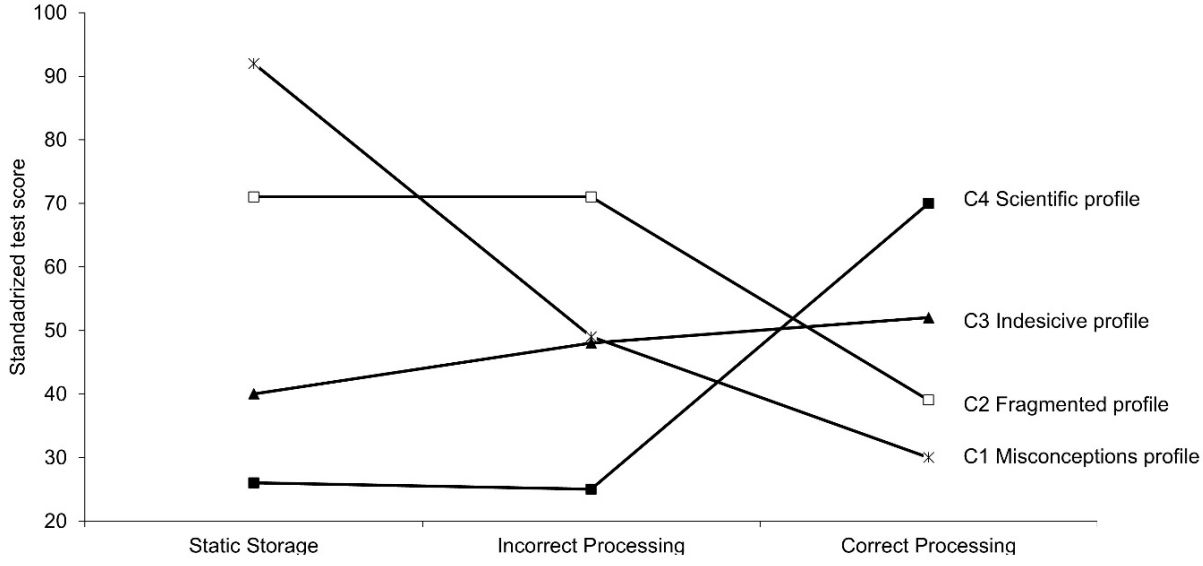

Figure 9. Diagram of the four knowledge profiles.

7.4.3 Changing Profile Frequencies on the Sample Level over Time

The proportions of the latent profiles changed substantially from T1 to T4 (see in Table 8), except for the misconceptions profile, which was shown only by a small proportion of the sample (4-7%) at all four measurement points. In line with hypothesis H3b, the frequency of the fragmented profile decreased from 56 to 17%, but did not approach 0, thus indicating that some Psychology students still had fragmented knowledge about the memory concept at the end of their fourth semester. The proportion of the indecisive profile increased from T1 to T2 and stayed stable to T3, before finally dropping back to the initial level. The frequency of the scientific profile increased sharply from 0% at T1 to 37% at T2 and remained high afterwards. These findings are in line with our second hypothesis (H2b), that there is an overall trend towards a scientific understanding of human memory.

7.4.4 Strengths of the Predictive Relations between Knowledge Profiles

In order to test for possible predictive relationships between the knowledge profiles of T1 to T4, we generated three frequency tables (T1-T2, T2-T3, and T3-T4). We used the profile proportions and the transition probabilities based on the estimated model to compute profile frequencies. Then we build cross tabulations, including the profile frequencies at one point in time (rows) to predict the profile frequencies at the next point in time (columns). For the predictive value of T1 for T2, the chi-square test indicated a highly significant and strong positive relation, $\chi^2 = 174.37$, $df = 6$, $p < .001$, Cramer's V = .798. The relations were about equally strong for the other points of time, $\chi^2 = 313.32$, $df = 9$, $p < .001$, Cramer's V = .876 (T2xT3), and $\chi^2 = 266.08$, $df = 9$, $p < .001$, Cramer's V = .808 (T3xT4). Thus, the knowledge profiles allow predictions about students' future knowledge profiles at a later point in time. Only 23% of the participants did not change their knowledge profile over the course of the study. For the remaining 77% of the sample, information about the profile at one point in time helped to predict changes to another profile. In the following section, we take a closer look at the participants' individual transition paths.

7.4.5 Latent Transition Paths

MPlus reports overall sample latent transition probabilities based on the estimated model across time (for the latent transition probabilities in the current study, see Appendix G). These probabilities are used to calculate latent transition paths. In a model with four profiles at four measurement points, there are $4^4 = 256$ possible latent transition paths between the profiles across time. In line with our Hypothesis 1, the participants were on few of these possible paths. There were only six paths taken by at least 5% of the sample (see Figure 10), and 81% of the sample was on these six paths. Approximately 17% of the sample was on another eight paths which were taken by at least one person of the sample. The remaining 3% were on paths that were estimated to be taken by less than one person of the sample by MPlus. The fact that a high proportion of the sample followed a small number of developmental paths shows that knowledge development was highly systematic in our sample.

We interpreted the transition paths based on between which profiles the transitions occurred and assigned labels to them (see Table 9). The first path P1 was labelled improving correct knowledge because participants on this path displayed an indecisive profile at T1 and transitioned to the scientific profile at T2, where they stayed throughout the course of the study. The second path P2 was called decreasing fragmentation because individuals on this path transitioned from the fragmented profile to the indecisive profile which reflects a development towards more integrated pieces of knowledge. Path P3 were called enduring fragmentation, as participants on this path stayed with the fragmented profile throughout the whole study. Similarly, participants on P4 stayed with the indecisive profile, thus we labelled the path enduring indecision. The paths P5 and P6 both included two transitions, and students on these paths started with the fragmented profile, transitioned to the indecisive profile, and finally moved to the scientific profile. Students on P5, transitioned to the scientific profile between T2 and T3, whereas students on P6 transitioned to the scientific profile between T3 and T4. We labelled the paths slowly evolving scientific concepts and quickly evolving

scientific concept because these transitions represent a gradual decrease in the static storage misconception while processing statements are more and more endorsed.

As can be seen in Figure 10, on the six most frequent transition paths, the participants either stayed in their respective latent profile over time or transitioned from lower-ranking profiles (e.g., the indecisive profile) to higher-ranking profiles (e.g., the scientific profile). There were no transitions in the opposite direction. Between T1 and T2, the majority (58%) of the sample transitioned from a lower-ranking to a higher-ranking profile. Conversely, between T2, T3, and T4 the majority (72%) of the sample stayed in their respective latent profile. This mirrors the fact that most participants attended a lecture on human memory during the first semester of their program.

Table 9

Pathways of Conceptual Change (Model-Estimated Latent Transition Paths) in the Sample

Path	Label	Knowledge Profile				Proportion of the sample (%)	
		T1	T2	T3	T4	Alone	Accumulated
P1	Increasing correct knowledge	C3	C4	C4	C4	25	25
P2	Decreasing Fragmentation	C2	C3	C3	C3	22	47
P3	Enduring Fragmentation	C2	C2	C2	C2	15	62
P4	Enduring Indecisiveness	C3	C3	C3	C3	8	70
P5	Slowly Evolving Scientific Concepts	C2	C3	C3	C4	6	76
P6	Quickly Evolving Scientific Concepts	C2	C3	C4	C4	5	81
P7-P14	Various Paths found for at least one person	Various profiles	Various profiles	Various profiles	Various profiles	17	97
P15-P256	Various Paths not found for at least one person	Various profiles	Various profiles	Various profiles	Various profiles	3	100

Note. C1 = misconceptions profile; C2 = fragmented profile; C3 = indecisive profile; C4 = scientific profile

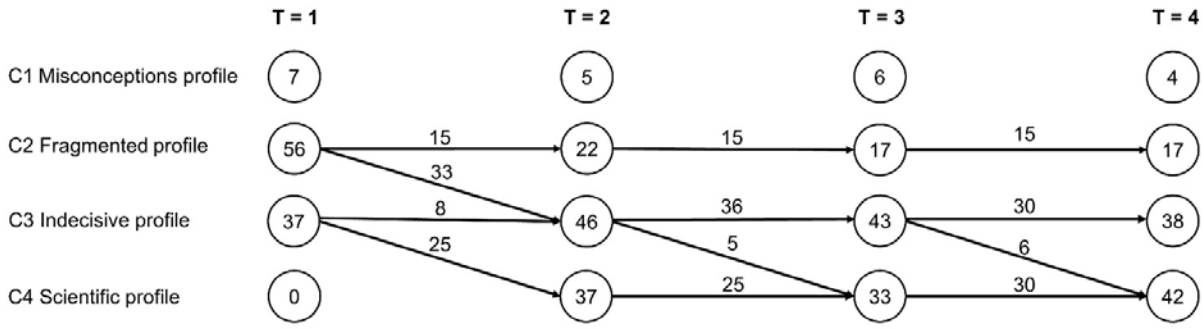

Figure 10. Diagram of the six transition paths taken at least by 5% of the sample. All numbers are percentages of the sample. The numbers in circles refer to the whole sample (100% in total at each wave). The numbers next to the arrows refer to the proportion of the sample on the most common six transitions paths (81% of the sample in total).

7.4.6 Associations without Outside Criteria

To test for relations between the knowledge profiles and outside criteria, we exported the list of participants' most likely profile memberships to SPSS. To test statistically for associations between profile memberships at T3 and T4, IMTS scores, and grades, IMTS and grades were recoded, so that higher scores indicated better achievement or better grades. As the number of students differed strongly between the four knowledge profiles, we used non-parametric tests to conduct the analyses. Parametric tests, such as the analysis of variance (ANOVA) have been shown to be robust against some violations of assumptions (Schmider, Ziegler, Danay, Beyer, & Bühner, 2010). However, multiple problems and unequal group sizes lead to serious constraints to the robustness of parametric tests (Lix & Keselman, 1998). As a non-parametric alternative to one-way ANOVA we used the Kruskal-Wallis H-test. To follow up on the findings, we used the Jonckheere-Terpstra test (Jonckheere, 1954), which allows to test whether the medians of the groups are ordered in a meaningful way. In SPSS, the Jonckheere-Terpstra test investigates whether the medians ascend or descend in a pre-defined order. We specified the independent variable in the hypothesized developmental rank order of knowledge profiles, that is, from a misconceptions profile, over fragmented and

indecisive profiles to a scientific profile. Table 10 shows the recoded IMTS scores and grades by most likely profile memberships.

As hypothesized (H4), the latent profiles at T3 differed significantly in their IMTS scores as indicated by the Kruskal-Wallis *H*-test, $H(3) = 25.845$, $p < .001$. Jonckheere's test was significant, $J = 3112.500$, $z = 4.546$, $r = .422$, indicating that participants' knowledge profiles predicted their achievement on the IMTS in the assumed developmental rank order from the misconceptions profile to the scientific profile. Thus, persons with a better understanding of memory in general are also more skeptical about the credibility of autobiographical memories. This is in line with a view of human memory as involving the active processing and reconstruction of information.

Persons differing in their knowledge profiles at T4 also differed systematically in their grades on the human memory course, $H(3) = 16.122$, $p = .001$. Jonckheere's test indicated that there was a significant trend in the data, indicating that grades improved from the misconceptions profile, over the fragmented profile and the indecisive profile, to the scientific profile, $J = 2714.500$, $z = 3.547$, $r = .335$, which supports our hypothesis H5.

Table 10

Descriptive Statistics for IMTS Scores and Grades by Most Likely Profile Memberships

		IMTS T3			Grade T4		
	Label	*M*	*SD*	*n*	*M*	*SD*	*n*
C1	Misconceptions profile	25.5	10.4	7	2.00	.628	5
C2	Fragmented profile	18.4	10.0	20	1.22	1.06	19
C3	Indecisive profile	30.3	8.43	53	1.97	.860	43
C4	Scientific profile	33.7	8.27	36	2.28	.686	46

Note. IMTS scores and grades are coded so that higher scores reflect higher competence.

7.5 Discussion

7.5.1 Knowledge Profiles

Our aim was to investigate whether conceptual change is still a relevant learning process in higher education students, therefore we conducted the first latent profile transition analysis in this population. In line with our Hypothesis 1, we found that there were individual differences in Psychology students' knowledge about human memory. Our analyses show that these individual differences of the 137 learners can be described in terms of only four knowledge profiles.

The misconceptions profile, in which students agreed mostly with static storage statements, was found in less than 10% of the sample at all four measurement points, respectively. This does not mean that the students in the sample started the study with perfect prior knowledge, though. The most frequent profile at the beginning of the first semester was the fragmented profile characterized by strong agreement with both static storage statements and incorrect processing statements. These two kinds of statements are mutually incompatible, because either knowledge is static or is processed in memory. We therefore interpreted students' agreement with both kinds of statements as indicating fragmented knowledge, in which the students hold both kinds of beliefs but do not understand their inter-relation. In line with Hypotheses 3a and 3b, the fragmented profile occurred frequently at the beginning of the study, decreased in its frequency over time, but was still found for 17% of the sample at the end of the study.

Many other students seemed to be aware of their lack of a complete understanding of memory, as the indecisive profile was the second most frequent at the first and the last measurement point and the most frequent on the second and third measurement point. Students with this profile tended to choose the middle category when evaluating statements, thus, neither agreeing nor disagreeing with them. Thus, not all students in our sample held deeply entrenched misconceptions. Instead, some students were aware of their lack of

understanding. Students with the scientific profile agreed mostly with correct processing statements. They did not only understand that memory actively constructs and re-constructs information, but also in which situations chunking, interference, or source monitoring are the most relevant memory processes.

The strongly decreasing sample proportions (from 56% at T1 to 17 % at T4) for the fragmented profile and the strongly increasing sample proportions for the scientific profile (from 0 to 42%) demonstrated a trend towards more scientifically correct and also more integrated knowledge over the course of the four semesters (Hypotheses 2b and 3b). These findings with respect to the knowledge profiles and their changing frequencies over time are generally well in line with the findings obtained by Michael Schneider and Hardy (2013) as well as Edelsbrunner, Schalk, Schumacher, and Stern (2015) with elementary school children. These two studies likewise found learners with fragmented knowledge and learners with integrated knowledge in their sample. Also, they found that fragmented knowledge decreased in its frequency in the sample but remained in some learners, and found an overall trend from misconceptions and fragmented knowledge to more integrated and correct knowledge. The main difference between the previous and the present findings is that the most frequent profile before learning was the misconceptions profile in the elementary school children and the fragmented profile in the university students. Future studies will have to investigate whether this is a general difference between the two age groups. Possibly, the longer learning history of the university students led to a great base of badly integrated knowledge in comparison to the school children.

7.5.2 Transition Paths

Students' knowledge changes over the four measurement points could be described in terms of few transition paths between the four knowledge profiles. There were only six paths taken by at least 5% of the sample, respectively, and a total of 81% of the sample was on these paths. The remaining part of the sample was on one of eight additional paths. The

participants on the six most common paths either stayed in their respective latent profile or moved to a higher ranking profile, as shown in Figure 10. Thus, there was a clear developmental ordering of the profiles from the misconceptions profile, over the fragmented profile and the indecisive profile to the scientific profile. The knowledge profiles differed in how strongly participants agreed with statements about static storage concepts, incorrect processing concepts, or correct processing concepts of human memory. We thus interpret transitions between these knowledge profiles as evidence of conceptual change. The involved profiles demonstrate the strength of conceptual change underlying the transition: Transitions between more similar profiles (e.g., the misconceptions profile and the indecisive profile) arguably require less knowledge restructuring than transitions between dissimilar profiles (e.g., the misconceptions profile and the scientific profile).

The highly systematic pathways imply that the learners' knowledge at each point in time is a good predictor of the learners' knowledge at later points in time. Indeed, we found significant and very strong relations between the knowledge profiles at different points in time. The fact that on the six most common paths more than half of the participants transitioned from one profile to another, that is, restructured their knowledge shows that conceptual change is a relevant learning mechanism in higher education and in the domain of Psychology. Zero percent of the participants had the scientific profile at the start of the study. The 42 % of the participants having this profile at the end of the study all transitioned there from lower-ranking knowledge profiles over time. This demonstrates that conceptual change does not only happen but is a central learning mechanism in acquiring academic concepts in higher education.

The findings also demonstrate the importance of knowledge fragmentation and integration for understanding and predicting conceptual change in higher education. As predicted, knowledge fragmentation was frequent at the start of the study, stayed constant on

one path (the enduring fragmentation path), and decreased in its frequency on other paths (e.g., the decreasing fragmentation path and the slowly evolving scientific concepts path).

The findings demonstrate systematic associations of the profiles and pathways with domain-specific instruction, a finding also observed in the study by Michael Schneider and Hardy (2013). Overall, 58% of the persons on the most frequent six paths transitioned to a different profile and 33% integrated previously fragmented knowledge during the first semester, during which most participants also attended a lecture about human memory. In contrast, only 11% of the participants transitioned to a different profile during the second, third, and fourth semester together, and no participants integrated their knowledge during that time.

7.5.3 Validity of the Model Results and Limitations

At least five findings indicated that the model results reflect systematic and meaningful individual differences rather than random measurement error. First, the transition paths and the predictive relations between the knowledge profiles show a high degree of systematic organization in the results both for individual differences at each point in time and intra-individual differences over time. Second, the basic pattern of results in the current study conceptually replicates the key findings of Michael Schneider and Hardy (2013), such as the small numbers of profiles and pathways, the co-existence of fragmented and integrated knowledge in the sample, and general trends towards more integrated and correct knowledge. Third, transitions to higher-ranking profiles were more frequent during domain-specific instruction (i.e., the lecture on human memory during the first semester) than at other times. Fourth, the latent profiles differed in their mean rank-ordered grades, with students in higher-ranking profiles also showing better grades. Though evidence from analyses with most likely profile variables can be somewhat limited when entropy is not perfect, we believe that in our study, in which entropy was only slightly smaller than .90, such bias was only a marginal problem. Therefore, the findings support the criterion validity of the latent profile variables

and demonstrate the grade-relevance of academic concepts and conceptual change in higher education. Finally, students in higher-ranking profiles also showed better performance on the implicit memory theory scale, indicating that a better understanding of memory as involving the construction and re-construction of information went along with a greater skepticism with respect to the contents of autobiographical memories.

We expect that the findings of our study are highly generalizable, as the interpretation our knowledge profiles is in line with previous studies conducted outside higher education (Edelsbrunner et al., 2015; Hardy et al., 2006; McMullen et al., 2015). Our findings are somewhat limited as the identification of knowledge profiles in latent transition analysis is exploratory. There are no definite standards in choosing the number of profiles. In latent transition analysis, the number of estimated parameter rises exponentially with the number of profiles assumed. With 63 free parameters in our model with four profiles at four measurement points and 137 participants, our design is at the margin of being underpowered. Our sample is restricted to psychology students that represent a highly selected, presumably high-achieving subpopulation. Future studies should investigate whether students from other programs than psychology show similar knowledge profiles and developmental patterns. Another limitation is the use of a newly constructed test in our study. Future studies will have to further examine the validity and reliability of the test and its sensitivity to knowledge changes over time.

7.5.4 Theoretical, Methodological, and Practical Implications

Our findings have theoretical, methodological, and practical implications. On the theoretical level, the results demonstrate that conceptual change still takes place in higher education students and is vital for acquiring an understanding of academic concepts. Conceptual change, its measurement, its prevalence, its dynamics over time, and its stimulation by learning environments are well investigated for school students but almost not examined at all for students in higher education. Our present findings suggest that future

studies should systematically investigate parallels and possible differences between conceptual change in K-12 school and in higher education. This will lead to a better understanding of which processes are specific to conceptual change in general and which processes only occur in specific age groups. Further, future studies should include motivational and affective covariates into latent transition models in order to investigate non-cognitive determinants of conceptual change. The small number of profiles and pathways indicated that individual differences matter in conceptual change, but that the number of possible individual differences is quite limited. This suggests that the learners' idiosyncratic knowledge construction processes are constrained by cognitive and environmental variables which guide these processes along a few developmental trajectories.

On the methodological level, our findings indicate that latent profile transition analyses are useful tools for modeling longitudinal changes in multidimensional knowledge structures in general and for modeling conceptual change in specific. Only very few latent profile transition analyses have been published so far (see Hickendorff et al., in press). The convergence of findings from earlier studies with school children and the current study with Psychology students indicates a high degree of replicability, stability, and generalizability of the findings across studies, age groups and content domains. Latent profile transition analyses have also proven to be effective data reduction techniques, because they help to describe individual differences in large samples in terms of just a few profiles and transition paths.

On the practical level, findings from studies like ours can inform higher education teaching. Teachers in higher education might often lack awareness that students enter their programs with prior beliefs and misconceptions, because higher education learning has long been neglected in research on conceptual change. The profiles and pathways identified in the current study can help teachers to better identify students' current knowledge and to predict students' future pathways of learning. Instruction can thus be tailored to fit subpopulations of students differing in their prior knowledge or development. The current study as well as the

study by Michael Schneider and Hardy (2013) also found evidence of systematic relations between instruction and the knowledge profiles and pathways. Future studies should investigate these interactions in greater detail.

8. General Discussion

This dissertation aimed to investigate environmental and learner-related determinants of knowledge acquisition. This goal was attained by means of two meta-analyses and one latent transition analysis. The presented studies provide six key insights regarding the relevance of environmental and individual factors accounting for knowledge acquisition and possible underlying mechanisms. The six key insights are discussed in the upcoming paragraphs, followed by methodological and practical implications, four recommendations for future research, and a final conclusion.

8.1 Key Insights from the Studies

The results from the three studies can be summarized in the following six key insights. First, tES significantly contributes to the enhancement of cognitive learning outcomes, such as test performance and knowledge acquisition. Second, results give implications that specific physiological processes of neuroplasticity underlie knowledge acquisition. Third, prior knowledge significantly contributes to later cognitive learning outcomes including knowledge and academic achievement. Fourth, results provide evidence for the existence of both, a Matthew effect as well as a compensatory effect, depending on characteristics of the instruction employed. Fifth, results give implications that cognitive processes underlie the acquisition of knowledge and that these can be accounted for in the instructional design. Sixth, despite that learners differ in their prior knowledge, only a limited amount of learning pathways become evident across learners when acquiring conceptual knowledge.

8.1.1 Key Insight 1

The first study demonstrated the effectiveness of tES, by combining results from single studies meta-analytically. Overall, the effect was significantly positive and medium in magnitude with Cohen's $d = 0.343$ (Jacob Cohen, 1988). The effect was consistently found in randomized controlled trials, and across a variety of stimulation methods, stimulation areas, measures, stimulus and task types, and domains. The results give evidence that the signal was

even strong enough to show significant results through the noise due to the heterogeneity in competence measures and activated brain regions.

8.1.2 Key Insight 2

The results from Study 1 give implications that specific physiological processes of neuroplasticity underlie knowledge acquisition. When tES was applied in a learning phase, where participants had to encode new information, a much stronger effect (d = 0.712) resulted compared to when tES was applied during test performance (d = 0.207), where participants had to recall information. The results may be explained by processes of synaptic plasticity (i.e., changes within GABAerig interneurons or NMDA receptor-dependent glutameric interneurons; Liebetanz, Nitsche, Tergau, & Paulus, 2002; Nitsche et al., 2003; Nitsche et al., 2005) and nonsynaptic plasticity (i.e., changes in membrane excitability; Nitsche et al., 2003), both online and offline stimulation (Stagg et al., 2009). This highlights the relevance of different neurophysiological processes underlying knowledge acquisition.

8.1.3 Key Insight 3

The results from Study 2 demonstrate that individual differences in prior knowledge are highly predictive for later knowledge and achievement at posttest with r = .525 what is equivalent to a Cohen's d of 1.23. The effect was consistently shown for different educational levels, domains, knowledge types, instructional methods, and measures. The effect of prior knowledge can be interpreted as causal rather than correlational, as a significant and positive effect was found for studies experimentally inducing differences in prior knowledge. Further, adjusting for intelligence did not significantly reduce the magnitude of the effect size, highlighting the robustness of the found effect.

8.1.4 Key Insight 4

The meta-analytic results of Study 2 provide evidence for both, a Matthew effect as well as a compensatory effect. Some studies demonstrated a positive correlation between prior knowledge and knowledge gains, indicating that high prior knowledge is associated with

stronger growth in knowledge as described by the Matthew effect. Others found a negative relationship between prior knowledge and knowledge gains, providing evidence for a compensatory effect. Whether a Matthew or compensatory effect became evident was dependent on the cognitive demands of the intervention.

8.1.5 Key Insight 5

The moderator analyses performed in Study 2 provide implications that cognitive processes underlying knowledge acquisition are a relevant factor which can be accounted for by instructional design. The relationship between prior knowledge and posttest knowledge was smaller in magnitude when instruction had higher cognitive demands as compared to lower cognitive demands. A similar pattern was found for the relationship between prior knowledge and knowledge gains. When cognitive demands of the instruction were high, learners with high prior knowledge benefitted more as compared to learners with low prior knowledge (Matthew effect). The reverse was found for instruction with lower cognitive demands (compensatory effect). Here, low prior knowledge learners benefitted more, as compared to their more knowledgeable counterparts. A moderating effect of the cognitive demands of the intervention was found for all learning outcome measures (posttest knowledge and knowledge gains), indicating the robustness of this effect. Thus, applying instructional design principles in teaching can significantly determine whether high or low knowledgeable students experience changes in knowledge, respectively.

8.1.6 Key Insight 6

Study 3 investigated the development of conceptual knowledge longitudinally, accounting for different qualities of knowledge by means of latent profile transition analyses. Despite the fact that learners significantly differed in their prior knowledge at pretest, the path of transitions between knowledge states over time was limited, indicating specific and well-ordered developmental pathways. The results demonstrate that even at the level of higher education, students still incorporate intuitive and fragmented prior knowledge, not compatible

with scientific models, which highlights the importance of individual differences in prior knowledge. More specifically, prior knowledge as a relevant determinant for knowledge acquisition can not only be described in terms of quantity of knowledge but also in terms of quality of knowledge such as fragmentation vs. integration, and incorrectness vs. correctness.

8.2 Methodological Implications

The results of the presented studies provide methodological implications to be considered and addressed by future studies, namely the operationalization of learning outcomes, the use of latent transition analyses in educational research, and the deployment of meta-analysis in educational contexts, along with the application of appropriate statistical procedures. The three implications are addressed in the following three paragraphs, respectively.

8.2.1 Operationalization of Learning Outcomes

When investigating the impact of individual differences in prior knowledge on effects of an intervention or learning in general, the study groups usually significantly differ in their knowledge before entering the intervention (for example high and low prior knowledge groups). Differences in the posttest may be hard to interpret as they may either be a result of the group differences at the pretest, an intervention (if applied), or the respective interaction. Therefore, it is more appropriate to consider the pretest values in the form of gain scores, especially when analyzing the effect of individual differences in prior knowledge on subsequent learning. Gain scores can be calculated in two different ways: as absolute gain scores and normalized gain scores. Absolute gain scores have the methodological disadvantage that learners with high prior knowledge have less room for learning gains as compared to learners with low prior knowledge. Thus, absolute learning gain scores might underestimate the actual gains of learners with high prior knowledge due to ceiling effects, and might lead to an underestimation of the influence of prior knowledge on knowledge acquisition. Normalized gain scores account for this problem by dividing the absolute

learning gain by the maximal possible learning gain. The results from Study 2 indicate significant differences in the effect of prior knowledge on the three outcome measures (correlation between pretest and posttest values vs. absolute gains vs. normalized gains), highlighting the relevance of accurately distinguishing between learning outcomes (posttest scores) and knowledge gains (absolute or normalized).

Overall, considering information from a pretest in the form of gain scores is accompanied with several benefits and some disadvantages. The benefits include a decrease in estimates of error variance, resulting in increased power to be able to account for individual differences between groups, and to allow for a precise interpretation of effect sizes due to an objective measure of gains in knowledge (e.g., Allison, 1990; Williams & Zimmerman, 1996). Disadvantages consists of limits in the application, such as the same parallel test, or Rasch measurement that needs to be employed at pre- and posttest (Andrich, 1988; Kline, 1986). Further, possible bias may occur due floor or ceiling effects, however, these can be eluded by the application of adaptive testing approaches (Rock, 2007).

A variety of statistical approaches exist to investigate changes in knowledge, for example, the use of difference scores (e.g., the calculation of normalized gain scores based on group means) regression based methods (e.g., the application of an ANCOVA with the pre-test as a covariate), or structural equation measures (e.g., growth curve analysis). It is likely that the application of gain scores has rarely been employed as researchers assumed that gain scores are unreliable (e.g., Cribbie & Jamieson, 2004; Cronbach & Furby, 1970). As a result, studies in the educational context commonly employ an ANCOVA to control for differences at pretest (which has been one of the most common reasons of exclusion of single studies in Study 2). However, there is evidence that this approach is not suited when differences between groups exist at pretest (see e.g., Gregory A. Miller & Chapman, 2001). More recent work has shown that reliability is not a serious problem when applying gain scores (e.g., Llabre, Spitzer, Saab, Ironson, & Schneiderman, 1991; Williams & Zimmerman, 1996). Out

of all 223 studies included in the meta-analysis presented in Study 2, only 14 studies reported absolute gain scores, and only three studies reported normalized gain scores as outcomes. This indicates that gain scores are still used infrequently, even in studies aiming to investigate the effects of individual differences in prior knowledge on knowledge acquisition. It is an open question regarding why so many studies failed to report gain scores although the data would have allowed such calculations. Moreover, more systematic analyses of the different gain score calculations (for example, see Vickers, 2001), appropriate measures that allow such calculations (for example, see Rock, 2007), and statistical analyses suited for gain scores (for example, see Bast & Reitsma, 1997; Bonate, 2000; Cribbie & Jamieson, 2004; T. E. Duncan, Duncan, Strycker, Li, & Alpert, 1999) are necessary to allow researchers to employ the adequate application of gain scores in their empirical work.

8.2.2 Latent Transition Analyses in Educational Research

Latent transition analyses are useful tools for modeling longitudinal change in multidimensional knowledge structures. Study 3 provides an example of how latent transition analysis can be used for modeling conceptual change and including knowledge measures that relate to specific qualities of knowledge (i.e., fragmented knowledge, misconceptions). Applying a statistical approach that can quantify differences and changes in knowledge quantity and knowledge quality is fundamental to comprehensively understand the principles of knowledge acquisition. Therefore, employing latent transition analysis can significantly contribute to a better understanding of longitudinal changes in multidimensional variables such as knowledge. Further, Study 3 demonstrated that latent transition analysis is an effective data reduction technique, as it allows for a description of individual differences in large samples in terms of few profiles and transition paths. This essentially facilitates a better understanding of the complex principles of knowledge acquisition. Another advantage of latent transition analysis is the possibility to predict the transition by third variables, for example intelligence or academic self-efficacy in educational research. This may further

provide valuable insights regarding relevant variables influential to changes in knowledge structures.

There are several limitations when applying latent transition analyses. First, difficulties in inference and estimation may occur, especially with small samples (Chung, Lanza, & Loken, 2008) or even medium-sized samples (Lanza & Collins, 2008). The development of some methodological aspects of latent transition analysis, for example power analyses to estimate the needed sample size for a respective analysis, is still outstanding. Second, in some cases, specific conclusions and recommendations drawn from the results may be difficult due to complex data (e.g., Lanza, Patrick, & Maggs, 2010) or having small classes (Pan, Liu, Lau, & Luo, 2017), especially when analyzing transition probabilities.

Overall, latent transition analysis is seen as a suitable tool for identifying qualitative intra- and inter-individual differences in the context of learning (Hickendorff et al., in press) and to analyze the development of these differences over time. The application may lead to new insights and a deeper understanding of knowledge acquisition over time and is especially appropriate for medium to large samples.

8.2.3 Meta-Analyses in Educational Research

Meta-analyses have some advantages over narrative reports or systematic reviews of the literature, and are therefore widely employed, especially in the context of educational psychology. There exist over 1200 meta-analyses relating to achievement just in the educational field (Hattie, 2015), and many more will likely be conducted in the future. The method for conduction a meta-analysis has come to fruition over the last four decades. This is mostly due to the enormous number of quantitative findings from single studies, driving the need for aggregation and summarization, and the desire to identify definite conclusions through aggregation (e.g., Hattie, 2009; S. Higgins, 2016). Meta-analyses in educational context may for example investigate the effectiveness of a specific intervention or school-instruction, or the associations between two or more variables of interest. Some meta-analyses

compare one treatment or multiple treatment groups with a control group (e.g., Dochy, Segers, Van den Bossche, & Gijbels, 2003; Kluger & DeNisi, 1996; Nesbit & Adesope, 2006). Others look at the association between two variables in the form of a correlation, most commonly the relationship between a specific variable and academic achievement (e.g., Fan & Chen, 2001; Hyde, Fennema, & Lamon, 1990; Richardson, Abraham, & Bond, 2012; Michael Schneider et al., 2017). Fewer meta-analyses have investigated other (non-cognitive) outcomes instead of academic achievement (e.g., Nurmi, 2012; Tanner-Smith & Wilson, 2013). The results from meta-analyses on educational topics can lead to validation or development of theory, best-practice recommendations for educators, the provision of information for policy makers, or demonstration of a typical effect, also referred to as hinge point data (e.g., Greenhalgh, 1997; Hattie, 2009; Kluger & DeNisi, 1996).

While meta-analysis is commonly employed in educational research, there is still a need to more systematically analyze different statistical approaches employed in meta-analyses. This is especially applicable for the dependency of data and effect sizes. There are multiple ways to calculate effect sizes from raw data presented in the single studies, especially for effect sizes obtained from repeated measures designs (e.g., Lakens, 2013; Morris & DeShon, 2002). As demonstrated by simulation analysis, the calculation of the effect size from the tests statistics reported in the single studies likely leads to overestimation, when the correlation between the measures is not considered (Dunlap, Cortina, Vaslow, & Burke, 1996). Becoming apparent in Studies 1 and 2, single studies reporting data for repeated measures rarely, if at all, report the correlation between the measurement points. This lack of information might consequently lead to an overestimation of effect sizes obtained in meta-analyses including effect sizes of dependent measures. The same concern may apply to the dependency of effect sizes obtained from the same sample. In Studies 1 and 2, the problem of dependency across effect sizes was successfully addressed by means of robust variance estimation (Hedges et al., 2010; Tanner-Smith & Tipton, 2014). As this method is rather new,

the development of an adequate method to perform publication-bias analyses with dependent data it is still outstanding. Further development and extension of the method seem promising to allow more sophisticated analyses.

Summarizing, meta-analysis is a useful tool to combine existing research, extend knowledge and theories within a specific discipline, and to draw attention on existing strengths and weaknesses in the literature (Carter, Kofler, Forster, & McCullough, 2015; Hunter & Schmidt, 1996; van Elk et al., 2015). It is argued that results from meta-analysis should be interpreted with the same caution as results from single studies (Carter et al., 2017). In addition to meta-analysis, large-scale studies and preregistered replication efforts are needed to carefully re-examine methodology and theory (e.g., Hagger et al., 2016). Due to the great heterogeneity in meta-analytic strategies employed, developing and following rigorous guidelines when conducting a meta-analysis would be especially useful (e.g., Carter et al., 2017; J. P. T. Higgins & Green, 2008; Moher, Liberati, Tetzlaff, & Altman, 2009; Shea et al., 2009). Further, more simulation analyses are necessary to fully understand how different methods of effect size computation and aggregation influence the results of meta-analyses. The extensions of statistical methods seem promising to allow for an adequate analysis of dependent effect sizes in meta-analysis.

8.3 Practical Implications

Several practical implications can be concluded from the results of the three studies presented in this dissertation. From Study 1, it is clear that more studies are needed before tES can be applied in educational settings, as research included in the synthesis was conducted in laboratory settings and only with adult participants. Therefore, the application of tES in educational settings should be considered with high caution. tES can be easily employed, as demonstrated in a variety of tutorials on YouTube showing how to build simple devices from everyday objects. Thus, teachers and educators need to be informed about new technologies

available that may potentially support learning, their benefits, possible harm and impairments (e.g., Davis, 2014; Wurzman et al., 2016), as well as the current research results. This is especially necessary for topics that not only attract researchers, but also the public, as is apparent with tES (Cohen Kadosh et al., 2012; Dubljevic, Saigle, & Racine, 2014). The effects of tES have been found particularly when a learning intervention was implemented, demonstrating the independent relevance of teaching and instruction for knowledge acquisition. Further, a variety of more effective and cheaper instructional methods exist compared to implementing tES (for example, see Hattie, 2009), leading to the question of the cost-benefit ratio.

The results provide valuable implications for teaching practices. As demonstrated in Study 2, the cognitive demands of the instruction employed can either contribute to a Matthew effect (i.e., promoting knowledge acquisition for high prior knowledge students) or a compensatory effect (i.e., promoting knowledge acquisition for the low prior knowledge students). When the cognitive demands of a specific intervention or instruction are high, it is likely that the high prior knowledge students benefit more as compared to their less knowledgeable counterparts. The reverse occurs for instruction with lower cognitive demands. Therefore, teachers can employ instructional methods tailored to fit the subpopulations of students differing in their prior knowledge or development (e.g., Kalyuga et al., 2003; Mayer & Moreno, 2003). For example, this can be achieved with the use of short summaries (Mayer, Bove, Bryman, Mars, & Tapangco, 1996), the implementation of knowledge maps in teaching (O'Donnell, Dansereau, & Hall, 2002), or the realization of complex programs promoting classroom instruction (Hindman, Erhart, & Wasik, 2002). The results further highlight the relevance of misconceptions for conceptual change in subsequent learning. Thus, it seems useful that teachers account for this by addressing their students' common misconceptions and tailoring school instruction accordingly (McCuin, Hayhoe, & Hayhoe, 2014). Possible instructional methods to facilitate conceptual change include

plausibility reappraisals (Lombardi, Sinatra, & Nussbaum, 2013) or problem based learning (Loyens, Jones, Mikkers, & van Gog, 2015). While teachers are usually aware of students' misconceptions, they may have difficulties using their students' prior knowledge in an effective way to adjust their teaching methods accordingly (Cox, Steegen, & De Cock, 2016). Hence, teacher education should emphasize training teachers on skills related to instructional design, especially considering prior knowledge and individual learning pathways. This would allow teachers to adjust their teaching based on the students' individual differences.

8.4 Recommendations for Future Research on Knowledge Acquisition

The results from the presented studies provide a valuable starting point for future research, indicating that there are remaining gaps in the literature. Generally, future research on knowledge acquisition should be characterized by a more detailed approach in order to deeply understand the specific mechanisms underlying knowledge acquisition. Further, it seems promising to integrate the views and perspectives from different sub-disciplines within the field of psychology as well as across different disciplines in general, to allow the benefit in form of interdisciplinary research. In the following, four recommendations for future research are proposed and discussed.

8.4.1 Recommendation 1: Consider Domain-Specific Knowledge, Including Different Types and Qualities of Knowledge

Compared to academic achievement, the employment of domain-specific knowledge as the dependent or independent variable in research has the advantage of potentially providing deeper insights into the processes of knowledge acquisition. Academic achievement has been widely used as an outcome variable and proxy for learning success. More precisely, learning success can be described in terms of changes in knowledge within a specific domain, assessed with a single or multiple knowledge tests. Most commonly, researchers subsume variables such as knowledge, results from standardized achievement tests, or grades under the broader

category of academic achievement. However, it is relevant to distinguish between measures assessing domain specific knowledge from more curriculum-based criteria, such as grades, and cumulative measures such as GPA, which are more general (Steinmayr et al., 2014). Prior knowledge stored in long-term memory is directly linked to changes in knowledge, as proposed by information processing theory (Atkinson & Shiffrin, 1968; Mayer & Moreno, 2003; George A. Miller, 1956). Assessing knowledge as a dependent or independent variable can therefore provide a more detailed picture of the mechanisms and underlying processes. Including domain-specific knowledge measures instead of grades or GPA has the further benefit of providing the possibility to consider different types and qualities of knowledge. Most commonly, knowledge is assessed as the amount of correct knowledge. However, knowledge can differ among several dimensions of type and quality (de Jong & Ferguson-Hessler, 1996), which can highly contribute to or impact knowledge acquisition (e.g., Friege & Lind, 2006). For example, incorrect knowledge in form of misconceptions or fragmented knowledge has a robust influence on later knowledge acquisition (diSessa et al., 2004; Smith III. et al., 1993; Vosniadou & Skopeliti, 2017). Therefore, it is desirable to closely investigate different features of knowledge, develop reliable tests to assess these differences, and employ such tests in future studies.

8.4.2 Recommendation 2: Analyze Prior Knowledge Effects on Knowledge Gains and Transfer

Prior knowledge has been demonstrated to be similarly related to learning outcomes, including knowledge and achievement. Until now, only few studies have investigated the effect of prior knowledge on knowledge gains (i.e., knowledge acquisition) and virtually no studies exist that manipulated prior knowledge and investigated its effect on knowledge gains. Thus, it remains unclear whether the effect of knowledge on knowledge acquisition is causal in nature. This reveals a large deficiency in educational research, as it has often been assumed that prior knowledge is closely related to learning (e.g., Ausubel, 1968). Further, even less is

known about potential third variables moderating this relationship. Several variables have been identified and analyzed that contribute to the understanding of acquisition of knowledge and skills, such as interest (e.g., Reeve & Hakel, 2000), intelligence (e.g., Giofrè, Borella, & Mammarella, 2017; Pfost et al., 2012), socio-economic status (e.g., Chiu, Chow, & Joh, 2017), or home and learning-related environmental conditions (e.g., Duff et al., 2015; Kempe, Eriksson-Gustavsson, & Samuelsson, 2011; Otto & Kistner, 2017; Pfost et al., 2012). However, systematic investigations of third variables relevant for changes in knowledge and knowledge acquisition have just begun recently, leaving many questions open to answer.

Further, it is still unclear whether and to what extent transfer of knowledge to novel contexts occurs (Perkins & Salomon, 1992; D. L. Schwartz, Varma, & Martin, 2008). This existing gap in research may be a result of the complexity of the topic, as different dimensions along which transfer can occur have been specified (e.g., different contents and contexts of transfer; Barnett & Ceci, 2002). More so, this is seen as one of the most challenging, contentious, and important issues in psychology (Day & Goldstone, 2012, p. 1). While there are contradictory empirical findings on transfer across single studies, only a few meta-analytic summaries with correlational or domain-specific evidence exist (Study 2 of this dissertation; Blume, Ford, Baldwin, & Huang, 2010; Melby-Lervåg & Lervåg, 2011; Pai, Sears, & Maeda, 2015). A meta-analysis summarizing the existing body of research that has engaged with transfer is still outstanding. Consideration of differences in prior knowledge when investigating transfer seems promising, as prior knowledge is associated with an enhanced ability to take advantage of deep structural content (Schoenfeld & Herrmann, 1982; Weiser & Shertz, 1983; Wolpert, 1990). Further, this facilitates both reminding and mapping (Day & Goldstone, 2012; Novick, 1988), and fosters the ability to overcome limitations in working memory associated with relevant mechanisms underlying transfer (Halford, Wilson, & Phillips, 1998; Waltz, Lau, Grewal, & Holyoak, 2000).

8.4.3 Recommendation 3: Search for Processes that Explain the Association Between Prior Knowledge and Knowledge Acquisition

There exists comprehensive evidence that biological and cognitive processes underlie knowledge acquisition, and that these processes can be determined by both environmental and individual characteristics. While there is empirical confirmation for the positive association between prior knowledge and neurophysiological processes (e.g., Brod, Lindenberger, Wagner, & Shing, 2016; Brod et al., 2013; Chang, Kang, Ham, & Lee, 2016; Liu, Grady, & Moscovitch, 2017; Shing & Brod, 2016) as well as cognitive processes (Chuy, Alamargot, & Passerault, 2012; Crooks & Alibali, 2013; DeWitt et al., 2012; Park, Korbach, & Brunken, 2015), which are both subsequently associated with knowledge acquisition (e.g., Giofrè et al., 2017; Jaekel, Eryigit-Madzwamuse, & Wolke, 2016; Kepinska, de Rover, Caspers, & Schiller, 2017; Minichiello, 2009; Vandenbrouke, Verschueren, & Baeyens, 2017), this evidence has not yet be summarized in a meta-analysis. Such analyses would significantly contribute to an enhanced understanding of knowledge acquisition and its underlying processes, and would further allow the testing of moderating and mediating effects. While most studies have only investigated one path, i.e., the association between prior knowledge and intra-learner processes or the association between intra-learner processes and learning outcomes, less is known about the mutual associations between prior knowledge, intra-learner processes and learning outcomes (e.g., Hemmer & Persaud, 2014; Kalyuga, 2013; Wu & Zhao, 2017).

8.4.4 Recommendation 4: Take a Broader and More Integrative View of Knowledge Acquisition

Research on knowledge acquisition has mainly focused on the investigation of the effectiveness of specific interventions, the relationship between variables, or the existence of specific paradigms found in an educational context. Less literature exists pertaining to the integration of findings on a superordinate level, and only a few models and theories exist that

include different research fields of educational science (Biggs, 2003, p. 19; Gagné, 2004; Stern, 2017; Ullén et al., 2016). Most of these models focus on outcomes other than knowledge, such as expertise (Ullén et al., 2016) or competencies (Gagné, 2004). Therefore, it seems beneficial to integrate research on knowledge acquisition on a subordinate level by means of theoretical and empirical approaches. The framework presented in Study 2 is one theoretical approach to combine relevant variables for knowledge acquisition identified by different streams of educational research. Meta-analytic and second-order meta-analytic investigations relating to associations presented in the framework would greatly contribute to a better understanding of knowledge acquisition, its underlying principles, and determining variables. Further, it seems promising to promote inter-disciplinary research (Ledford, 2015; Science Europe, 2012), to link research fields to achieve a more comprehensive approach to address critical questions, enhance the robustness and validity of findings by use of different populations and approaches, and ultimately facilitate the application of knowledge in a specific area (e.g., Aboelela et al., 2007; F. Fisher et al., 2014; Tandon & Singh, 2016).

9. Conclusion

The studies included in the current dissertation contributed to enhancing the understanding of how environmental and learner-related variables determine knowledge acquisition. The results demonstrated that learning outcomes can be advanced by input from the environment, and are highly related to the students' level of prior knowledge. Further, the effects of environmental and learner-related variables interact and likely impact the biological and cognitive processes underlying learning. Research combining approaches from different disciplines such as neuroscience, cognitive science, and educational science seems promising to provide an overall picture of relevant determinants for knowledge acquisition. There is a necessity to apply appropriate and recent statistical approaches to fully understand the underlying processes of knowledge acquisition, for example the development and

employment of a rigid definition of learning outcomes and its operationalization, the utilization of latent transition analyses to investigate longitudinal changes in knowledge structures, or the use of meta-analyses to aggregate existing research. The results presented in this doctoral thesis demonstrate that knowledge acquisition is a dynamic process involving biological and cognitive mechanisms and being moderated by the environment and the single learner. Accounting for different variables associated with knowledge acquisition and academic achievement leads to a better understanding of the underlying complex processes and greater possibilities of the advancements of learning.

Appendix

Appendix A: Included Studies (Study 1)

Balconi, M., & Vitaloni, S. (2014). Dorsolateral pFC and the representation of the incorrect use of an object: the transcranial dirrect current stimulation effect on N400 for visual and linguistic stimuli. *Journal of Cognitive Neuroscience, 26*(2), 305-318. doi:10.1162/jocn_a_00500

Beharelle, A. R., Polanía, R., Hare, T. A., & Ruff, C. C. (2015). Transcranial stimulation over frontopolar cortex elucidates the choice attributes and neural mechanisms used to resolve exploration-exploitation trade-offs. *The Journal of Neuroscience, 35*(43), 14544-14556. doi:10.1533/JNEUROSCI.2322-15.2015

Cappelletti, M., Gessaroli, E., Hithersay, R., Mitolo, M., Didino, D., Kanai, R., . . . Walsh, V. (2013). Transfer of cognitive training across magnitude dimensions achieved with concurrent brain stimulation of the parietal lobe. *The Journal of Neuroscience, 33*(37), 14899–14907. doi:10.1523/JNEUROSCI.1692-13.2013

Cappelletti, M., Pikkat, H., Upstill, E., Speekenbrink, M., & Walsh, V. (2015). Learning to integrate versus inhibiting information is modulated by age. *The Journal of Neuroscience, 35*(5), 2213-2225. doi:10.1523/JNEUROSCI.1018-14.2015

Fertonani, A., Brambilla, M., Cotelli, M., & Miniussi, C. (2014). The timing of cognitive plasticity in physiological aging: a tDCS study of naming. *Frontiers in Aging Neuroscience, 6,* Article 131. doi:10.3389/fnagi.2014.00131

Fertonani, A., Rosini, S., Cotelli, M., Rossini, P. M., & Miniussi, C. (2010). Naming facilitation induced by transcranial direct current stimulation. *Behavioural Brain Research, 208,* 311–318. doi:10.1016/j.bbr.2009.10.030

Fiori, V., Coccia, M., Marinelli, C. V., Becchi, V., Bonifazi, S., Veravolo, M. G., . . . Marangolo, P. (2011). Transcranial direct current stimulation improves word retrieval

in healthy and nonfluent aphasic subjects. *Journal of Cognitive Neuroscience, 23*(9), 2309–2323. doi:10.1162/jocn.2010.21579

Flöel, A., Rösser, N., Michka, O., Knecht, S., & Breitenstein, C. (2008). Noninvasive brain stimulation improves language learning. *Journal of Cognitive Neuroscience, 20*(8), 1415–1422. doi:10.1162/jocn.2008.20098

Friederici, A. D., Mueller, J. L., Sehm, B., & Ragert, P. (2013). Language learning without control: the role of the PFC. *Journal of Cognitive Neuroscience, 25*(5), 814–821. doi:10.1162/jocn_a_00350

Grabner, R. H., Rütsche, B., Ruff, C. C., & Hauser, T. U. (2013). *Transcranial direct current stimulation to modulate learning of novel arithmetic procedures*. Unpublished manuscript.

Grabner, R. H., Rütsche, B., Ruff, C. C., & Hauser, T. U. (2015). Transcranial direct current stimulation of the posterior parietal cortex modulates arithmetic learning. *European Journal of Neuroscience, 42*, 1667–1674. doi:10.1111/ejn.12947

Hauser, T. U., Rotzer, S., Garbner, R. H., Mérillat, S., & Jäncke, L. (2013). Enhancing performance in numerical magnitude processing andmental arithmetic using transcranial Direct Current Stimulation (tDCS). *frontiers in Human Neuroscience, 7*(Article 244). doi:10.3389/fnhum.2013.00244

Henseler, I., Mädebach, A., Kotz, S. A., & Jescheniak, J. D. (2014). Modulating brain mechanisms resolving lexico-semantic interference during word production: a transcranial direct current stimulation study. *Journal of Cognitive Neuroscience, 26*(7), 1403–1417. doi:10.1162/jocn_a_00572

Houser, R., Thoma, S., Fonseca, D., O'Conner, E., & Stanton, M. (2015). Enhancing statistical calculation with transcranial direct current stimulation (tDCS) to the left intra-parietal sulcus (IPS). *Trends in Neuroscience and Education, 4*(4), 98-101. doi:10.1016/j.tine.2015.07.002

Hussey, E. K., Ward, N., Christianson, K., & Kramer, A. F. (2015). Language and memory improvements following tDCS of left lateral prefrontal cortex. *PLOS ONE, 10*(11), e0141417. doi:10.1371/journal.pone.0141417

Ihara, A. S., Mimura, T., Soshi, T., Yorfuji, S., Hirata, M., Goto, t., . . . Fujimaki, N. (2015). Facilitated lexical ambiguity processing by transcranial direct current stimulation over the left inferior frontal cortex. *Journal of Cognitive Neuroscience, 27*(1), 26–34. doi:10.1162/jocn_a_00703

Javadi, A. H., Brunec, I. K., Walsh, V., Penny, W. D., & Spiers, H. J. (2014). Transcranial electrical brain stimulation modulates neuronal tuning curves in perception of numerosity and duration. *NeuroImage, 102*, 451–457. doi:10.1016/j.neuroimage.2014.08.016

Klein, E., Mann, A., Huber, S., Bloechle, J., Willmes, K., Karim, A. A., . . . Moeller, K. (2013). Bilateral bi-cephalic tdcs with two active electrodes of the same polarity modulates bilateral cognitive processes differentially. *PLOS ONE, 8*(8), e71607. doi:10.1371/journal.pone.0071607

Li, L. M., Leech, R., Scott, G., Malhotra, P., Seemungal, B., & Sharp, D. J. (2015). The effect of oppositional parietal transcranial direct current stimulation on lateralized brain functions. *European Journal of Neuroscience, 42*, 2904–2914. doi:10.1111/ejn.13086

Looi, C. Y., Duta, M., Brem, A.-K., Huber, S., Nuerk, H.-C., & Cohen Kadosh, R. (2016). Combining brain stimulation and video game to promote long-term transfer of learning and cognitive enhancement. *Nature, 6*, 22003. doi:10.1038/srep22003

Meinzer, M., Jähningen, S., Copeland, D. A., Darkow, R., Grittner, U., Avirame, K., . . . Flöel, A. (2014). Transcranial direct current stimulation over multiple days improves learning and maintenance of a novel vocabulary. *Cortex, 50*, 137-147. doi:10.1016/j.cortex.2013.07.013

Morales-Quezada, L., Cosmo, C., Carvalho, S., Leite, J., Castillo-Saavedra, L., Rozisky, J. R., & Fregni, F. (2015). Cognitive effects and autonomic responses to transcranial pulsed current stimulation. *Exp Brain Res, 233,* 701–709. doi:10.1007/s00221-014-4147-y

Nozari, N., Arnold, J. E., & Thompson-Schill, S. L. (2014). The effects of anodal stimulation of the left prefrontal cortex on sentence production. *Brain Stimul., 7*(6), 784-792. doi:10.1016/j.brs.2014.07.035.

Penolazzi, B., Pastore, M., & Mondini, S. (2013). Electrode montage dependent effects of transcranial direct current stimulation on semantic fluenc. *Behavioural Brain Research, 248,* 129– 135. doi:10.1016/j.bbr.2013.04.007

Peretz, Y., & Lavidor, M. (2013). Enhancing lexical ambiguity resolution by brain polarization of the right posterior superior temporal sulcus. *Cortex, 49*(4), 1056-1062. doi:10.1016/j.cortex.2012.03.015

Pope, P. A., Brenton, J. W., & Miall, R. C. (2015). Task-specific facilitation of cognition by anodal transcranial direct current stimulation of the prefrontal cortex. *Cerebral Cortex, 25*, 4551–4558. doi:10.1093/cercor/bhv094

Popescu, T., Krause, B., Terhune, D. B., Twose, O., Page, T., Humphreys, G., & Cohen Kadosh, R. (2016). Transcranial random noise stimulation mitigates increased difficulty in an arithmetic learning task. *Neuropsychologia, 81,* 255–264. doi:10.1016/j.neuropsychologia.2015.12.028

Price, A. R., Peelle, J. E., Bonner, M. F., Grossman, M., & Hamilton, R. H. (2016). Causal evidence for a mechanism of semantic integration in the angular gyrus as revealed by high-definition transcranial direct current stimulation. *The Journal of Neuroscience, 36*(13), 3829–3838. doi:10.1523/JNEUROSCI.3120-15.2016

Ross, L. A., McCoy, D., Coslett, H. B., Olson, I. R., & Wolk, D. A. (2011). Improved proper name recall in aging after electrical stimulation of the anterior temporal lobes. *Frontiers inAging Neuroscience, 3*(Article 16). doi:10.3389/fnagi.2011.00016

Rütsche, B., Hauser, T. U., Jäncke, L., & Garbner, R. H. (2015). When problem size matters: differential effects of brain stimulation on arithmetic problem solving and neural oscillations. *PLOS ONE, 10*(3). doi:10.1371/journal.pone.0120665

Sahlem, G. L., Badran, B. W., Halford, J. J., Williams, N. R., Korte, J. E., Leslie, K., . . . George, M. S. (2015). Oscillating square wave Transcranial Direct Current Stimulation (tDCS) delivered during slow wave sleep does not improve declarative memory more than sham: A randomized sham controlled crossover study. *Brain Stimul., 8*(3), 528–534. doi:10.1016/j.brs.2015.01.414.

Sarkar, A., Dowker, A., & Cohen Kadosh, R. (2014). Cognitive enhancement or cognitive cost: trait-specific outcomes of brain stimulation in the case of mathematics anxiety. *The Journal of Neuroscience, 34*(50), 16605–16610. doi:10.1523/JNEUROSCI.3129-14.2014

Snowball, A., Tachtsidis, I., Popescu, T., Thompson, J., Delazer, M., Zamarian, L., . . . Cohen Kadosh, R. (2013). Long-term enhancement of brain function and cognition using cognitive training and brain stimulation. *Current Biology, 23*, 987-992. doi:10.1016/j.cub.2013.04.045

Sparing, R., Dafotakis, M., Meister, I. G., Thirugnanasambandam, N., & Finke, G. R. (2008). Enhancing language performance with non-invasive brain stimulation—A transcranial direct current stimulation study in healthy humans. *Neuropsychologia, 46*, 261–268. doi:10.1016/j.neuropsychologia.2007.07.009

Vannorsdall, T. D., Schretlen, D. J., Andrejczuk, M., Ledoux, K., Bosley, L. V., R.Weaver, J., . . . Gordon, B. (2012). Altering automatic verbal processes with transcranial direct current stimulation. *Frontiers in Psychiatry, 3*, Article 73. doi:10.3389/fpsyt.2012.00073

Appendix B: Excluded Studies (Study 1)

Andrade, A. C., Magnavita, G. M., Allegro, J. V. B. N., Neto, C. E. B. P., Lucena, R. d. C. S., & Fregni, F. (2014). Feasibility of transcranial dirrect current stimulation use in children aged 5 to 12 years. *Journal of Child Neurology, 29*(10), 1360-1365. doi:10.1177/0883073813503710

Artemenko, C., Moeller, K., Huber, S., & Klein, E. (2015). Diffential influences of unilateral tDCS over the intraparietal cortex on numerical cognition. *Frontiers in Human Neuroscience, 9*(Article 110). doi:10.3389/fnhum.2015.00110

Clemens, B., Jung, S., Zvyagintsev, K., Domahs, F., & Willmes, K. (2013). Modulating arithmetic fact retrieval: A single-blind, sham-controlled tDCS study with repeated fMRI measurements. *Neuropsychologia, 51*, 1279–1286. doi:10.1016/j.neuropsychologia.2013.03.023

Hecht, D., Walsh, V., & Lavidor, M. (2010). Transcranial direct current Stimulation facilitates decision making in a probabilistic guessing task. *The Journal of Neuroscience, 30*(12), 4241– 4245. doi:10.1523/JNEUROSCI.2924-09.2010

Iuculano, T. (2014). Preliminary evidence for performance enhancement following parietal lobe stimulation in Developmental Dyscalculia. *Frontiers in Human Neuroscience, 8*(38). doi:doi: 10.3389/fnhum.2014.00038

Kotilainen, T., Lehto, S. M., & Wikgren, J. (2015). Effect of transcranial direct current stimulation on semantic discrimination eyeblink conditioning. *Behavioural Brain Research, 292*, 142-146. doi:10.1016/j.bbr.2015.06.021

Wolkenstein, L., Zeiller, M., Kanske, P., & Plewnia, C. (2014). Induction of a depression-like negativity bias by cathodal transcranial direct current stimulation. *Cortex, 59*, 103-112. doi:10.1016/j.cortex.2014.07.011

Appendix C: Included Studies (Study 2)

Aarnoutse, C., Van Leeuwe, J., & Verhoeven, L. (2005). Early Literacy From a Longitudinal Perspective. *Educational Research and Evaluation, 11*(3), 253-275. doi:10.1080/08993400500101054

Acuna, S. R., Rodicio, H. G., & Sanchez, E. (2011). Fostering active processing of instructional explanations of learners with high and low prior knowledge. *European Journal of Psychology of Education, 26*(4), 435-452. doi:10.1007/s10212-010-0049-y

Alao, S., & Guthrie, J. T. (1999). Predicting conceptual understanding with cognitive and motivational variables. *The Journal of Educational Research, 92*(4), 243-254.

Alonzo, C. N., Yeomans-Maldonado, G., Murphy, K. A., & Bevens, B. (2016). Predicting second grade listening comprehension using prekindergarten measures. *Topics in Language Disorders, 36*(4), 312-333. doi:10.1097/TLD.0000000000000102

Amadieu, F., Tricot, A., & Marine, C. (2009). Prior knowledge in learning from a non-linear electronic document: Disorientation and coherence of the reading sequences. *Computers in Human Behavior, 25*(2), 381-388. doi:10.1016/j.chb.2008.12.017

Amadieu, F., Tricot, A., & Marine, C. (2010). Interaction between prior knowledge and concept-map structure on hypertext comprehension, coherence of reading orders and disorientation. *Interacting with computers, 22*(2), 88-97. doi:10.1016/j.intcom.2009.07.001

Amadieu, F., Van Gog, T., Paas, F., Tricot, A., & Marine, C. (2009). Effects of prior knowledge and concept-map structure on disorientation, cognitive load, and learning. *Learning and Instruction, 19*(5), 376-386. doi:10.1016/j.learninstruc.2009.02.005

Aram, D., Korat, O., & Hassunah-Arafat, S. (2013). The contribution of early home literacy activities to first grade reading and writing achievements in Arabic. *Reading and Writing, 26*(9), 1517-1536. doi:10.1007/s11145-013-9430-y

Armand, F. (2001). Learning from expository texts; Effects of the interaction of prior knowledge and text structure on responses to different question types. *European Journal of Psychology of Education, 16*(1), 67-86.

Aunola, K., Leskinen, E., Lerkkanen, M.-K., & Nurmi, J.-E. (2004). Developmental Dynamics of Math Performance From Preschool to Grade 2. *Journal of Educational Psychology*, *96*(4), 699-713. doi:10.1037/0022-0663.96.4.699

Bailey, D. H., Siegler, R. S., & Geary, D. C. (2014). Early predictors of middle school fraction knowledge. *Developmental Science, 17*(5), 775-785. doi:10.1111/desc.12155

Bailey, D. H., Watts, T. W., Littlefield, A. K., & Geary, D. C. (2014). State and trait effects on individual differences in children's mathematical development. *Psychological Science, 25*(1). doi:10.1177/0956797614547539

Balch, W. R. (2015). Related General-Vocabulary Knowledge Transfers to Learning Technical Terms. *Teaching of Psychology*, *42*(3), 254-259. doi:10.1177/0098628315589495

Bergey, B. W., Cromley, J. G., Kirchgessner, M. L., & Newcombe, N. S. (2015). Using diagrams versus text for spaced restudy: Effects on learning in 10th grade biology classes. *British Journal of Educational Psychology*, *85*(1), 59-74. doi:10.1111/bjep.12062

Blayney, P. J., Kalyuga, S., & Sweller, J. (2015). Using Cognitive Load Theory to Tailor Instruction to Levels of Accounting Students' Expertise. *Educational Technology & Society*, *18*(4), 199-210.

Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Development, 79*(4), 1016-1031. doi: 10.1111/j.1467-8624.2008.01173.x

Braasch, J. L. G., & Goldman, S. R. (2010). The role of prior knowledge in learning from analogies in science texts. *Discourse Processes, 47*(6), 447-479. doi:10.1080/01638530903420960

Braten, I., Anmarkrud, O., Brandmo, C., & Stromso, H. I. (2014). Developing and testing a model of direct and indirect relationships between individual differences, processing, and multiple-text comprehension. *Learning and Instruction, 30*, 9-24. doi:10.1016/j.learninstruc.2013.11.002

Bråten, I., Johansen, R.-P., & Strømsø, H. I. (2017). Effects of different ways of introducing a reading task on intrinsic motivation and comprehension. *Journal of Research in Reading, 40*(1), 17-36. doi:10.1111/1467-9817.12053

Braten, I., & Stromso, H. I. (2011). Measuring strategic processing when students read multiple texts. *Metacognition and Learning, 6*(2), 111-130. doi:10.1007/s11409-011-9075-7

Bryant, P., Nunes, T., & Barros, R. (2014). The connection between children's knowledge and use of grapho-phonic and morphemic units in written text and their learning at school. *British Journal of Educational Psychology, 84*(2), 211-225. doi:1 0.1 1 1 l/bjep.12030

Bulu, S. T., & Pedersen, S. (2012). Supporting problem-solving performance in a hypermedia learning environment: The role of students' prior knowledge and metacognitive skills. *Computers in Human Behavior, 28*(4), 1162-1169. doi:10.1016/j.chb.2012.01.026

Burgess, S. R. (2002). The influence of speech perception, oral language ability, the home literacy environment, and pre-reading knowledge on the growth of phonological sensitivity: A one-year longitudinal investigation. *Reading and Writing, 15*(7-8), 709-737.

Butcher, K. R., & Sumner, T. (2011). How does prior knowledge impact students' online learning behaviors? . *International Journal of Cyber Behavior, Psychology and Learning, 1*(4), 1-18. doi:10.4018/ijcbpl.2011100101

Byrnes, J. P. (1992). The conceptual basis of procedural learning. *Cognitive Development, 7*(2), 235-257.

Callahan, D., & Drum, P. A. (1984). Reading ability and prior knowledge as predictors of eleven and twelve year olds' text comprehension. *Reading Psychology, 5*(1-2), 145-154.

Caravolas, M., Hulme, C., & Snowling, M. (2001). The foundations of spelling ability: Evidence from a 3-year longitudinal study. *Journal of Memory and Language, 45*(4), 751-774.

Carroll, J. M. (2004). Letter knowledge precipitates phoneme segmentation, but not phoneme invariance. *Journal of Research in Reading, 27*(3), 212-225.

Carroll, J. M., Snowling, M., Stevenson, J., & Hulme, C. (2003). The development of phonological awareness in preschool children. *Developmental Psychology, 39*(5), 913-923. doi:10.1037/0012-1649.39.5.913

Carson, K., Boustead, T., & Gillon, G. (2015). Content validity to support the use of a computer-based phonological awareness screening and monitoring assessment (Com-PASMA) in the classroom. *International Journal of Speech-Language Pathology, 17*(5), 500-510. doi:10.3109/17549507.2015.1016107

Cavallo, A. M. L., & Schafer, L. E. (1994). Relationships between students' meaningful learning orientation and their understanding of genetics topics. *Journal of Research in Science Teaching, 31*(4), 393-418.

Chambers, S. K., & Andre, T. (1997). Gender, prior knowledge, interest, and experience in electricity and conceptual change text manipulations in learning about direct current. *Journal of Research in Science Teaching, 34*(2), 107-123. doi:10.1002/(SICI)1098-2736(199702)34:2<107::AID-TEA2>3.0.CO;2-X

Chandran, S., Treagust, D. F., & Tobin, K. (1987). The role of cognitive factors in chemistry achievement. *Journal of Research in Science Teaching, 24*(2), 145-160. doi:10.1002/tea.3660240207

Chen, C.-H., & Wu, I.-C. (2012). The interplay between cognitive and motivational variables in a supportive online learning system for secondary physical education. *Computers & Education, 58*(1), 542-550. doi:10.1016/j.compedu.2011.09.012

Cheng, Y., Zhang, J., Wu, X., Liu, H., & Li, H. (2016). Cross-lagged relationships between morphological awareness and reading comprehension among chinese children. *Frontiers in Psychology, 7*, 1379. doi:10.3389/fpsyg.2016.01379

Chliounaki, K., & Bryant, P. (2007). How children learn about morphological spelling rules. *Child Development, 78*(4), 1360-1373.

Cho, J.-R., Chiu, M. M., & McBride-Chang, C. (2011). Morphological awareness, phonological awareness, and literacy development in Korean and English: A 2-year longitudinal study. *Scientific Studies of Reading, 15*(5), 383-408. doi:10.1080/10888438.2010.487143

Chu, F. W., vanMarle, K., & Geary, D. C. (2016). Predicting children's reading and mathematics achievement from early quantitative knowledge and domain-general cognitive abilities. *Frontiers in Psychology, 7*, 775. doi:10.3389/fpsyg.2016.00775

Coiro, J. (2011). Predicting reading comprehension on the internet: Contributions of offline reading skills, online reading skills, and prior knowledge. *Journal of Literacy Research, 43*(4), 352-392. doi:10.1177/1086296X11421979

Connor, C., Rice, D. C., Canto, A. I., Southerland, S. A., Underwood, P., Kaya, S., . . . Morrison, F. J. (2012). Child characteristics by science instruction interactions in second and third grade and their relation to students' content-area knowledge, vocabulary, and reading skill gains. *The Elementary School Journal, 113*(1), 52-75.

Conrad, N. J., & Deacon, S. H. (2016). Children's orthographic knowledge and their word reading skill: testing bidirectional relations. *Scientific Studies of Reading, 20*(4), 339-347. doi:10.1080/10888438.2016.1183128

Crawford, J., Dancer, J., & Pittenger, J. (1986). Initial performance level on a speechreading task as related to subsequent improvement after shortterm training. *The Volta Review, 88*(2), 101-105.

Crevecoeur, Y. C., Coyne, M. D., & McCoach, D. B. (2014). English language learners and English-only learners' response to direct vocabulary instruction. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 30*(1), 51-78. doi:10.1080/10573569.2013.758943

Dahlstrom, O. (2012). Learning during a collaborative final exam. *Education Research and Evaluation, 18*(4), 321-332.

Dandache, S., Wouters, J., & Ghesquière, P. (2014). Development of reading and phonological skills of children at family risk for dyslexia: A longitudinal analysis from kindergarten to sixth grade. *Dyslexia, 20,* 305-329. doi:10.1002/dys.1482

Darabi, A., Arrastia-Lloyd, M. C., Nelson, D. W., Liang, X., & Farrell, J. (2015). Learning how the electron transport chain works: independent and interactive effects of instructional strategies and learners' characteristics. *Advances in Health Sciences Education, 20*(5), 1135-1148. doi:10.1007/s10459-015-9592-2

Davis, M. H., & Guthrie, J. T. (2015). Measuring reading comprehension of content area texts using an assessment of knowledge organization. *The Journal of Educational Research, 108*(2), 148-164.

Diakidoy, I.-A. N., Mouskounti, T., Fella, A., & Ioannides, C. (2016). Comprehension processes and outcomes with refutation and expository texts and their contribution to learning. *Learning and Instruction, 41,* 60-69. doi:10.1016/j.learninstruc.2015.10.002

Diakidoy, I.-A. N., Mouskounti, T., & Ioannides, C. (2011). Comprehension and learning from refutation and expository texts. *Reading Research Quarterly, 46*(1), 22-38. doi:10.1598/RRQ.46.l.2

Diana, E. M., & Webb, J. M. (1997). Using geographic maps in classrooms: The conjoint influence of individual differences and dual coding on learning facts. *Learning and Individual Differences, 9*(3), 195-214. doi:10.1016/S1041-6080(97)90006-6

Dittman, C. K. (2016). The impact of early classroom inattention on phonological processing and word-reading development. *Journal of Attention Disorders, 20*(8), 653-664. doi:10.1177/1087054713478979

Dresel, M., Ziegler, A., Broome, P., & Heller, K. A. (1998). Gender differences in science education: The double-edged role of prior knowledge in physics. *Roeper Review, 21*(2), 101-106. doi:10.1080/02783199809553939

Ercetin, G. (2010). Effects of topic interest and prior knowledge on text recall and annotation use in reading a hypermedia text in the L2. *ReCALL: Journal of Eurocall, 22*(2), 228-246. doi:10.1017/S0958344010000091

Fan, Y.-C., Wang, T.-H., & Wang, K.-H. (2011). A web-based model for developing assessment literacy of secondary in-service teachers. *Computers & Education, 57*(2), 1727-1740. doi:10.1016/j.compedu.2011.03.006

Fesel, S. S., Segers, E., Clariana, R. B., & Verhoeven, L. (2015). Quality of children's knowledge representations in digital text comprehension: Evidence from pathfinder networks. *Computers in Human Behavior, 49*, 135-146. doi:10.1016/j.chb.2015.01.014

Florez, I. R. (2011). Case-based instruction in early childhood teacher preparation: Does it work? *Journal of Early Childhood Teacher Education, 32*(2), 118-134. doi:10.1080/10901027.2011.572227

Florit, E., Roch, M., & Levorato, M. C. (2014). Listening text comprehension in preschoolers: A longitudinal study on the role of semantic components. *Reading and Writing, 27*(5), 793-817. doi:10.1007/s11145-013-9464-1

Fricke, S., Szczerbinski, M., Fox-Boyer, A., & Stackhouse, J. (2016). Preschool predictors of early literacy acquisition in German-speaking children. *Reading Research Quarterly, 51*(1), 29-53. doi:10.1002/rrq.116

Fricke, S., Szczerbinski, M., Stackhouse, J., & Fox-Boyer, A. V. (2008). Predicting individual differences in early literacy acquisition in German: The role of speech and language processing skills and letter knowledge. *Written Language and Literacy, 11*(2), 103-146. doi:10.1075/wll.11.2.02fri

Furnes, B., & Samuelsson, S. (2009). Preschool cognitive and language skills predicting kindergarten and grade 1 reading and spelling: A cross-linguistic comparison. *Journal of Research in Reading, 32*(3), 275-292. doi:JO.l l l l!j.1467-9817.2009.01393.x

Fyfe, E. R., & Rittle-Johnson, B. (2016). Feedback both helps and hinders learning: The causal role of prior knowledge. *Journal of Educational Psychology, 108*(1), 82-97. doi:10.1037/edu0000053

Fyfe, E. R., Rittle-Johnson, B., & DeCaro, M. S. (2012). The effects of feedback during exploratory mathematics problem solving: Prior knowledge matters. *Journal of Educational Psychology, 104*(4), 1094-1108. doi:10.1037/a0028389

Garcia-Madruga, J. A., & Fernandez-Corte, T. (1999). Learning complex historical knowledge at high school: The role of working memory. *Analise Psicologica, 17*(2), 241-252.

García, J. R., Bustos, A., & Sánchez, E. (2015). The contribution of knowledge about anaphors, organisational signals and refutations to reading comprehension. *Journal of Research in Reading, 38*(4), 405-427. doi:10.1111/1467-9817.12021

Gay, G. (1986). Interaction of learner control and prior understanding in computer-assisted video instruction. *Journal of Educational Psychology, 78*(3), 225-227.

Georgiou, G. K., Torppa, M., Manolitsis, G., Lyytinen, H., & Parrila, R. (2012). Longitudinal predictors of reading and spelling across languages varying in orthographic consistency. *Reading and Writing, 25*(2), 321-346. doi:10.1007/s11145-010-9271-x

Gobel, S. M., Watson, S. E., Lervag, A., & Hulme, C. (2014). Children's arithmetic development: It is number knowledge, not the approximate number sense, that counts. *Psychological Science, 25*(3), 789-798. doi:10.1177/0956797613516471

Goodrich, J. M., Lonigan, C. J., & Farver, J. M. (2013). Do early literacy skills in children's first language promote development of skills in their second language? An experimental evaluation of transfer. *Journal of Educational Psychology, 105*(2), 414-426. doi:10.1037/a0031780

Greene, J. A., Costa, L.-J., Robertson, J., Pan, Y., & Deekens, V. M. (2010). Exploring relations among college students' prior knowledge, implicit theories of intelligence, and self-regulated learning in a hypermedia environment. *Computers & Education, 55*(3), 1027-1043. doi:10.1016/j.compedu.2010.04.013

Griggs, R. A., & Jackson, S. L. (1988). A reexamination of the relationship of high school psychology and natural science courses to performance in a college introductory psychology class. *Teaching of Psychology, 15*(3), 142-144.

Guo, J.-p., & Pang, M. F. (2011). Learning a mathematical concept from comparing examples: the importance of variation and prior knowledge. *European Journal of Psychology of Education, 26*(4), 495-525. doi:10.1007/s10212-011-0060-y

Haenggi, D., & Perfetti, C. A. (1992). Individual differences in reprocessing of text. *Journal of Educational Psychology, 84*(2), 182-192.

Haenggi, D., & Perfetti, C. A. (1994). Processing components of college-level reading comprehension. *Discourse Processes, 17*(1), 83-104.

Hailikari, T., Nevgi, A., & Komulainen, E. (2008). Academic self-beliefs and prior knowledge as predictors of student achievement in mathematics: A structural model. *Educational Psychology, 28*(1), 59-71.

Hall, V. C., & Edmondson, B. (1992). Relative importance of aptitude and prior domain knowledge on immediate and delayed posttests. *Journal of Educational Psychology, 84*(2), 219-223.

Hambrick, D. Z. (2003). Why are some people more knowledgeable than others? A longitudinal study of knowledge acquisition. *Memory & Cognition, 31*(6), 902-917. doi:10.3758/BF03196444

Heemsoth, T., & Heinze, A. (2014). The impact of incorrect examples on learning fractions: A field experiment with 6th grade students. *Instructional Science, 42*(4), 639-657. doi:10.1007/s11251-013-9302-5

Helle, L., Nivala, M., Kronqvist, P., Ericsson, K. A., & Lehtinen, E. (2010). Do prior knowledge, personality and visual perceptual ability predict student performance in microscopic pathology? *Medical Education, 44*(6), 621-629. doi:10.1111/j.1365-2923.2010.03625.x

Hollingsworth, P. M., & Reutzel, D. R. (1990). Prior knowledge, content-related attitude, reading comprehension: Testing Mathewson's affective model of reading. *The Journal of Educational Research, 83*(4), 194-199.

Homer, B. D., & Plass, J. L. (2010). Expertise reversal for iconic representations in science visualizations. *Instructional Science, 38*(3), 259-276. doi:10.1007/s11251-009-9108-7

Hou, H.-T., & Li, M.-C. (2014). Evaluating multiple aspects of a digital educational problem-solving-based adventure game. *Computers in Human Behavior, 30*, 29-38. doi:10.1016/j.chb.2013.07.052

Hulme, C., Bowyer-Crane, C., Carroll, J. M., Duff, F. J., & Snowling, M. (2012). The causal role of phoneme awareness and letter-sound knowledge in learning to read:

Combining intervention studies with mediation analyses. *Psychological Science, 23*(6), 572-577. doi:10.1177/0956797611435921

Huntsinger, C. S., Jose, P. E., & Luo, Z. (2016). Parental facilitation of early mathematics and reading skills andknowledge through encouragement of home-based activities. *Early Cildhood Research Quarterly, 37*, 1-15. doi:10.1016/j.ecresq.2016.02.005

Ise, E., Arnoldi, C. J., & Schulte-Korne, G. (2014). Development of orthographic knowledge in German-speaking children: A 2-year longitudinal study. *Journal of Research in Reading, 37*(3), 233-249. doi:10.1111/j.1467-9817.2012.01535.x

Jared, D., Cormier, P., Levy, B. A., & Wade-Woolley, L. (2011). Early predictors of biliteracy development in children in French immersion: A 4-year longitudinal study. *Journal of Educational Psychology, 103*(1), 119-139. doi:10.1037/a0021284

Johnson, A. M., Ozogul, G., & Reisslein, M. (2015). Supporting multimedia learning with visual signalling and animated pedagogical agent: Moderating effects of prior knowledge. *Journal of Computer Assisted Learning, 31*(2), 97-115. doi:10.1111/jcal.12078

Johnson, M. A., & Lawson, A. E. (1998). What are the relative effects of reasoning ability and prior knowledge on biology achievement in expository and inquiry classes? *Journal of Research in Science Teaching, 35*(1), 89-103. doi:10.1002/(SICI)1098-2736(199801)35:1<89::AID-TEA6>3.0.CO;2-J

Johnston, M. B., & Whitman, T. L. (1987). Enhancing math computation through variations in training format and instructional content. *Cognitive Therapy and Research, 11*(3), 381-397.

Johnston, P. H. (1984). Prior knowledge and reading comprehension test bias. *Reading Research Quarterly, 19*(2), 219-239. doi:10.2307/747364

Juul, H., Poulsen, M., & Elbro, C. (2014). Separating speed from accuracy in beginning reading development. *Journal of Educational Psychology, 106*(4), 1096-1106. doi:10.1037/a0037100

Kahn-Horwitz, J., Sparks, R. L., & Goldstein, Z. (2012). English as a foreign language spelling development: A longitudinal study. *Applied Psycholinguistics, 33*(2), 343-363.

Kalyuga, S. (2008). Relative effectiveness of animated and static diagrams: An effect of learner prior knowledge. *Computers in Human Behavior, 24*(3), 852-861. doi:10.1016/j.chb.2007.02.018

Kendeou, P., & van den Broek, P. (2007). The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition, 35*(7), 1567-1577.

Kerfoot, B. P., Shaffer, K., McMahon, G. T., Baker, H., Kirdar, J., Kanter, S., . . . Armstrong, E. G. (2011). Online "spaced education progress-testing" of students to confront two upcoming challenges to medical schools. *Academic Medicine, 86*(3), 300-306.

Kirby, J. R., & Boulter, D. R. (1999). Spatial ability and transformational geometry. *European Journal of Psychology of Education, 14*(2), 283-294.

Klauda, S. L., & Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational Psychology, 100*(2), 310-321. doi:10.1037/0022-0663.100.2.310

Klem, M., Hagtvet, B., Hulme, C., & Gustafsson, J.-E. (2016). Screening for language delay: Growth trajectories of language ability in low- and high-performing children. *Speech, Language, and Hearing Research, 59*, 1035-1045. doi:10.1044/2016_JSLHR-L-15-0289

Klem, M., Melby-Lervag, M., Hagtvet, B., Lyster, S.-A. H., Gustafsson, J.-E., & Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than

working memory limitations. *Developmental Science, 18*(1), 146-154. doi:10.1111/desc.12202

Kobayashi, K. (2009). The influence of topic knowledge, external strategy use, and college experience on students' comprehension of controversial texts. *Learning and Individual Differences, 19*(1), 130-134. doi:10.1016/jl inclif.2008.06.001

Kopcha, T. J., & Sullivan, H. (2008). Learner preferences and prior knowledge in learner-controlled computer-based instruction. *Educational Technology Research and Development, 56*(3), 265-286.

Kramarski, B., Weiss, I., & Sharon, S. (2013). Generic versus context-specific prompts for supporting self-regulation in mathematical problem solving among students with low or high prior knowledge. *Journal of Cognitive Education and Psychology, 12*(2), 197-214. doi:10. 1891/1945-8959.12 .2. 197

Krause, U.-M., Stark, R., & Mandl, H. (2009). The effects of cooperative learning and feedback on e-learning in statistics. *Learning and Instruction, 19*(2), 158-170. doi:10.1016/j.learninstruc.2008.03.003

Kruk, R. S., Mayer, J., & Funk, L. (2014). The predictive relations between non-alphanumeric rapid naming and growth in regular and irregular word decoding in at-risk readers. *Journal of Research in Reading, 37*(1), 17-35. doi:10.1111/jrir.12005

Kyun, S. A., & Lee, H. (2009). The effects of worked examples in computer-based instruction: Focus on the presentation format of worked examples and prior knowledge of learners. *Asia Pacific Education Review, 10*(4), 495-503. doi:10.1007/s12564-009-9044-x

Lambiotte, J. G., & Dansereau, D. F. (1992). Effects of knowledge maps and prior knowledge on recall of science lecture content. *Journal of Experimental Education, 60*(3), 189-201.

Landerl, K., & Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography: An 8-year follow-up. *Journal of Educational Psychology, 100*(1), 150-161. doi:10.1037/0022-0663.100.1.150

Lawson, A. E. (1983). Predicting science achievement: The role of developmental level, disembedding ability, mental capacity, prior knowledge, and beliefs. *Journal of Research in Science Teaching, 20*(2), 117-129. doi:10.1002/tea.3660200204

Lawson, A. E., & Worsnop, W. A. (1992). Learning about evolution and rejecting a belief in special creation: Effects of reflective reasoning skill, prior knowlege, prior belief and religious commitment. *Journal of Research in Science Teaching, 29*(2), 143-166. doi:10.1002/tea.3660290205

Le Bigot, L., & Rouet, J.-F. (2007). The impact of presentation format, task assignment, and prior knowledge on students' comprehension of multiple online documents. *Journal of Literacy Research, 39*(4), 445-470. doi:10.1080/10862960701675317

Lee, C.-Y., & Chen, M.-J. (2014). The impacts of virtual manipulatives and prior knowledge on geometry learning performance in junior high school. *Journal of Educational Computing Research, 50*(2), 179-201.

Lee, C. H., & Kalyuga, S. (2011). Effectiveness of on-screen pinyin in learning Chinese: An expertise reversal for multimedia redundancy effect. *Computers in Human Behavior, 27*(1), 11-15. doi:10.1016/j.chb.2010.05.005

Lee, K.-W. L., Goh, N.-K., Chia, L.-S., & Chin, C. (1996). Cognitive variables in problem solving in chemistry: A revisited study. *Science Education, 80*(6), 691-710.

Lee, Y., & Nelson, D. W. (2005). Viewing or visualising--Which concept map strategy works best on problem-solving performance? *British Journal of Educational Technology, 36*(2), 193-203. doi:10.1111/j.1467-8535.2005.00452.x

LeFevre, J.-A., Jimenez Lira, C., Sowinski, C., Cankaya, O., Kamawar, D., & Skwarchuk, S.-L. (2013). Charting the role of the number line in mathematical development. *Frontiers in Psychology, 4,* ArtID 641.

Lenzner, A., Schnotz, W., & Müller, A. (2013). The role of decorative pictures in learning. *Instructional Science, 41*(5), 811-831. doi:10.1007/s11251-012-9256-z

Lepola, J., Lynch, J., Kiuru, N., Laakkonen, E., & Niemi, P. (2016). Early oral language comprehension, task orientation, and foundational reading skills as predictors of grade 3 reading comprehension. *Reading Research Quarterly, 51*(4), 373-390. doi:10.1002/rrq.145

Lepola, J., Lynch, J., Laakkonen, E., Silven, M., & Niemi, P. (2012). The role of inference making and other language skills in the development of narrative listening comprehension in 4-6-year-old children. *Reading Research Quarterly, 47*(3), 259-282. doi:10.1002/rrq.020

Lepola, J., Poskiparta, E., Laakkonen, E., & Niemi, D. (2005). Development of and Relationship Between Phonological and Motivational Processes and Naming Speed in Predicting Word Recognition in Grade 1. *Scientific Studies of Reading, 9*(4), 367-399. doi:10.1207/s1532799xssr0904_3

Lerkkanen, M.-K., Rasku-Puttonen, H., Aunola, K., & Nurmi, J.-E. (2005). Mathematical performance predicts progress in reading comprehension among 7-year olds. *European Journal of Psychology of Education, 20*(2), 121-137. doi:10.1007/BF03173503

Lerner, M. D., & Lonigan, C. J. (2016). Bidirectional relations between phonological awareness and letter knowledge in preschool revisited: A growth curve analysis of the relation between two code-related skills. *Journal of Experimental Child Psychology, 144*(2016), 166-183. doi:10.1016/j.jecp.2015.09.023

Lervag, A., & Aukrust, V. G. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *Journal of Child Psychology and Psychiatry, 51*(5), 612-620. doi:10.1111/j.1469-7610.2009.02185.x.

Lervag, A., & Hulme, C. (2010). Predicting the growth of early spelling skills: Are there heterogeneous developmental trajectories? . *Scientific Studies of Reading, 14*(6), 485-513. doi:10.1080/10888431003623488

Leslie, L., & Cooper, P. (1993). Assessing the predictive validity of prior-knowledge assessmen. *National Reading Conference Yearbook, 42*, 93-99.

Leu, D. J., Forzani, E., Rhoads, C., Maykel, C., Kennedy, C., & Timbrell, N. (2015). The new literacies of online research and comprehension: Rethinking the reading achievement gap. *Reading Research Quarterly, 50*(1), 37-59.

Levin, I., Ravid, D., & Rapaport, S. (2001). Morphology and spelling among Hebrew-speaking children: From kindergarten to first grade. *Journal of Child Language, 28*(3), 741-772.

Li, T., McBride-Chang, C., Wong, A., & Shu, H. (2012). Longitudinal predictors of spelling and reading comprehension in Chinese as an L1 and English as an L2 in Hong Kong Chinese children. *Journal of Educational Psychology, 104*(2), 286-301.

Liebfreund, M. D. (2015). Success with informational text comprehension: An examination of underlying factors. *Reading Research Quarterly, 50*(4), 387-392. doi:10.1002/rrq.109

Liebfreund, M. D., & Conradi, K. (2016). Component skills affecting elementary students' informational text comprehension. *Reading and Writing, 29*(6), 1141-1160. doi:10.1007/s11145-016-9629-9

Lindsey, K. A., Manis, F. R., & Bailey, C. E. (2003). Prediction of first-grade reading in Spanish-speaking English-language learners. *Journal of Educational Psychology, 95*(3), 482-494. doi:10.1037/0022-0663.95.3.482

Linnenbrink-Garcia, L., Pugh, K. J., Koskey, K. L. K., & Stewart, V. C. (2012). Developing conceptual understanding of natural selection: The role of interest, efficacy, and basic prior knowledge. *The Journal of Experimental Education, 80*(1), 45-68. doi:10.1080/00220973.2011.559491

Liu, H.-C., & Chuang, H.-H. (2011). Investigation of the impact of two verbal instruction formats and prior knowledge on student learning in a simulation-based learning environment. *Interactive Learning Environments, 19*(4), 433-446. doi:10 .1080/10494820903356940

Liu, T.-C., Lin, Y.-C., & Paas, F. (2014). Effects of prior knowledge on learning from different compositions of representations in a mobile learning environment. *Computers & Education, 72*, 328-338. doi:10.1016/j.compedu.2013.10.019

Liu, Y., Lin, D., & Zhang, X. (2016). Morphological awareness longitudinally predicts counting ability in Chinese kindergarteners. *Learning and Individual Differences, 47*(2016), 215-221. doi:10.1016/j.lindif.2016.01.007

Lo, L.-y., Yeung, P.-s., Ho, C. S.-H., Chan, D. W.-o., & Chung, K. (2015). The role of stroke knowledge in reading and spelling in Chinese. *Journal of Research in Reading., 00*(00), 1-22. doi:10.1111/1467-9817.12058

Lonigan, C. J., Burgees, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental Psychology, 36*(5), 596-613.

McBride-Chang, C., & Ho, C. S.-H. (2005). Predictors of Beginning Reading in Chinese and English: A 2-Year Longitudinal Study of Chinese Kindergartners. *Scientific Studies of Reading, 9*(2), 117-144. doi:10.1207/s1532799xssr0902_2

McGeown, S. P., Johnston, R. S., & Medford, E. (2012). Reading instruction affects the cognitive skills supporting early reading development. *Learning and Individual Differences, 22*(3), 360-364. doi:10.1016/j.lindif.2012.01.012

Moos, D. C., & Azevedo, R. (2009). Self-efficacy and prior domain knowledge: To what extent does monitoring mediate their relationship with hypermedia learning? *Metacognition and Learning, 4*(3), 197-216. doi:10.1007/s11409-009-9045-5

Morris, D., Bloodgood, J. W., Lomax, R. G., & Perney, J. (2003). Developmental steps in learning to read: A longitudinal study in kindergarten and first grade. *Reading Research Quarterly, 38*(3), 302-328. doi:10.1598/RRQ.38.3.1

Mou, Y., Li, Y., Hoard, M. K., Nugent, L. D., Chu, F., W., Rouder, J. N., & Geary, D. C. (2016). Developmental foundations of children's fraction magnitude knowledge. *Cognitive Development, 39*(2016), 141-153. doi:10.1016/j.cogdev.2016.05.002

Müller-Kalthoff, T., & Möller, J. (2003). The Effects of Graphical Overviews, Prior Knowledge, and Self-Concept on Hypertext Disorientation and Learning Achievement. *Journal of Educational Multimedia and Hypermedia, 12*(2), 117-134.

Müller-Kalthoff, T., & Möller, J. (2006). Browsing while reading: Effects of instructional design and learners' prior knowledge. *ALT-J Association for Learning Technology Journal, 14*(2), 183-198. doi:10.1080/09687760600668602

Murphy, K. P., & Alexander, P. A. (2002). What counts? The predictive powers of subject-matter knowledge, strategic processing, and interest in domain-specific performance. *Journal of Experimental Education, 70*(3), 197-214. doi:10.1080/00220970209599506

Muter, V., Hulme, C., Snowling, M., & Stevenson, J. (2004). Phonemes, Rimes, Vocabulary, and Grammatical Skills as Foundations of Early Reading Development: Evidence From a Longitudinal Study. *Developmental Psychology, 40*(5), 665-681. doi:10.1037/0012-1649.40.5.665

Muter, V., Hulme, C., Snowling, M., & Taylor, S. (1997). Segmentation, not rhyming, predicts early progress in learning to read. *Journal of Experimental Child Psychology, 65*(3), 370-396. doi:10.1006/jecp.1996.2365

Muter, V., & Snowling, M. (1997). Grammar and phonology predict spelling in middle childhood. *Reading and Writing, 9*(5-6), 407-425. doi:10.1007/978-94-017-3054-9_6

Naslund, J. C., & Schneider, W. (1996). Kindergarten letter knowledge, phonological skills, and memory processes: Relative effects on early literacy. *Journal of Experimental Child Psychology, 62*(1), 30-59. doi:10.1006/jecp.1996.0021

Nguyen, T., Watts, T. W., Duncan, G. J., Clements, D. H., Sarama, J. S., Wolfe, C., & Spitler, M. E. (2016). Which preschool mathematics competencies are most predictive of fifth grade achievement? *Early Childhood Research Quarterly, 36*(2016), 550-560. doi:10.1016/j.ecresq.2016.02.003

Nielsen, A.-M. V., & Juul, H. (2016). Predictors of early versus later spelling development in Danish. *Reading and Writing, 29*(2), 245-266. doi:10.1007/s11145-015-9591-y

Nithart, C., Demont, E., Metz-Lutz, M.-N., Majerus, S., Poncelet, M., & Leybaert, J. (2011). Early contribution of phonological awareness and later influence of phonological memory throughout reading acquisition. *Journal of Research in Reading, 34*(3), 346-363. doi:10.1111/j.1467-9817.2009.01427.x

Nunes, T., Bryant, P., & Bindman, M. (2006). The effects of learning to spell on children's awareness of morphology. *Reading and Writing, 19*(7), 767-787. doi:10.1007/s11145-006-9025-y

O'Brien, L. M., Meszaros, B., & Pulliam, W. E. (1985). Effects of teachers' use of objectives on student achievement in social studies. *Theory and Research in Social Education, 13*(3), 57-65. doi:10.1080/00933104.1985.10505506

O'Brien, L. M., Paratore, J. R., Leighton, C. M., Cassano, C. M., Krol-Sinclair, B., & Green, J. G. (2014). Examining differential effects of a family literacy program on language and literacy growth of English language learners with varying vocabularies. *Journal of Literacy Research, 46*(3), 383-415. doi:10.1177/1086296X14552180

Oakhill, J. V., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading, 16*(2), 91-121. doi:10.1080/10888438.2010.529219

Ohst, A., Fondu, B., M.E. , Glogger, I., Nückles, M., & Renkl, A. (2014). Preparing learners with partly incorrect intuitive prior knowledge for learning. *Frontiers in Psychology, 5*(664), 1-11. doi:10.3389/fpsyg.2014.00664

Ohst, A., Glogger, I., Nückles, M., & Renkl, A. (2015). Helping preservice teachers with inaccurate and fragmentary prior knowledge to acquire conceptual understanding of psychological principles. *Psychology Learning & Teaching, 14*(1), 5-25. doi:10.1177/1475725714564925

Oortwijn, M., Boekaerts, M., & Vedder, P. (2008). The impact of the teacher's role and pupils' ethnicity and prior knowledge on pupils' performance and motivation to cooperate. *Instructional Science, 36*(3), 251-268. doi:10.1007/s11251-007-9032-7

Östergren, R., & Träff, U. (2013). Early number knowledge and cognitive ability affect early arithmetic ability. *Journal of Experimental Child Psychology, 115*(3), 405-421. doi:10.1016/j.jecp.2013.03.007

Pampush, J. D., & Petto, A. J. (2011). Familiarity with Latin and Greek anatomical terms and course performance in undergraduates. *Anatomical Sciences Education, 4*(1), 9-15. doi:10.1002/ase.189

Pan, J., McBride-Chang, C., Shu, H., Liu, H., Zhang, Y., & Li, H. (2011). What is in the naming? A 5-year longitudinal study of early rapid naming and phonological sensitivity in relation to subsequent reading skills in both native Chinese and English as a second language. *Journal of Educational Psychology, 103*(4), 897-908. doi:10.1037/a0024344

Park, S. I., Lee, G., & Kim, M. (2009). Do students benefit equally from interactive computer simulations regardless of prior knowledge levels? *Computers & Education, 52*(3), 649-655. doi:10.1016/j.compedu.2008.11.014

Passolunghi, M. C., Vercelloni, B., & Schadee, H. (2007). The precursors of mathematics learning: Working memory, phonological ability and numerical competence. *Cognitive Development, 22*(2), 165-184. doi:10.1016/j.cogdev.2006.09.001

Philips, B. M., Lonigan, C. J., & Wyatt, M. A. (2009). Predictive validity of the Get Ready to Read! screener: Concurrent and long-term relations with reading-related skills. *Journal of Learning Disabilities, 42*(2), 133-147. doi:10.1177/0022219408326209

Piasta, S. B., Petscher, Y., & Justice, L. M. (2012). How many letters should preschoolers in public programs know? The diagnostic efficiency of various preschool letter-naming benchmarks for predicting first-grade literacy achievement. *Journal of Educational Psychology, 104*(4), 945-958. doi:10.1037/a0027757

Piquard-Kipffer, A., & Sprenger-Charolles, L. (2013). Early predictors of future reading-skills: A follow-up of French-speaking children from the beginning of kindergarten to the end of the second grade (age 5 to 8). *L'Annee Psychologique, 113*(4), 491-521. doi:10.4074/S0003503313014012

Poe, M. D., Burchinal, M. R., & Roberts, J. E. (2004). Early language and the development of children's reading skills. *Journal of School Psychology, 42*(4), 315-332. doi:10.1016/j.jsp.2004.06.001

Priebe, S. J., Keenan, J. M., & Miller, A. C. (2012). How prior knowledge affects word identification and comprehension. *Reading and Writing, 25*(1), 131-149. doi:10.1007/s11145-010-9260-0

Quinn, J. M., Wagner, R. K., Petscher, Y., & Lopez, D. (2015). Developmental relations between vocabulary knowledge and reading comprehension: A latent change score modeling study. *Child Development, 86*(1), 159-175. doi:10.1111/cdev.12292

Randler, C., Ilg, A., & Kern, J. (2005). Cognitive and Emotional Evaluation of an Amphibian Conservation Program for Elementary School Students. *The Journal of Environmental Education, 37*(1), 43-52. doi:10.3200/JOEE.37.1.43-52

Recht, D. R., & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology, 80*(1), 16. doi:10.1037/0022-0663.80.1.16

Ree, M. J., Carretta, T. R., & Teachout, M. S. (1995). Role of ability and prior knowledge in complex training performance. *Journal of Applied Psychology, 80*(6), 721-730. doi:10.1037/0021-9010.80.6.721

Reisslein, J., Atkinson, R. K., Seeling, P., & Reisslein, M. (2006). Encountering the expertise reversal effect with a computer-based environment on electrical circuit analysis. *Learning and Instruction, 16*(2), 92-103. doi:10.1016/j.learninstruc.2006.02.008

Retelsdorf, J., & Köller, O. (2014). Reciprocal effects between reading comprehension and spelling. *Learning and Individual Differences, 30*, 77-83. doi:10.1016ij.lindif.2013.11 JJ07

Richland, L. E., & Burchinal, M. R. (2013). Early executive function predicts reasoning development. *Psychological Science, 24*(1), 87-92. doi:10.1177/0956797612450883

Rittle-Johnson, B., & Koedinger, K. (2009). Iterating between lessons on concepts and procedures can improve mathematics knowledge. *British Journal of Educational Psychology, 79*(3), 483-500. doi:10.1348/000709908X398106

Sadi, Ö., & Cakiroglu, J. (2014). Relations of cognitive and motivational variables with students' human circulatory system achievement in traditional and learning cycle classrooms. *Educational Sciences: Theory & Practice, 14*(5), 1997-2012. doi:10.12738/estp.2014.5.2037

Salmeron, L., Kintsch, W., & Canas, J. J. (2006). Reading strategies and prior knowledge in learning from hypertext. *Memory & Cognition, 34*(5), 1157-1171. doi:10.3758/BF03193262

Samuelstuen, M. S., & Braten, I. (2005). Decoding, knowledge, and strategies in comprehension of expository text. *Scandinavian Journal of Psychology, 46*(2), 107-117. doi:10.1111/j.1467-9450.2005.00441.x

Sanchez, M., Magnan, A., & Ecalle, J. (2012). Knowledge about word structure in beginning readers: What specific links are there with word reading and spelling? *European Journal of Psychology of Education, 27*(3), 299-317. doi:10.1007/s10212-011-0071-8

Sander, E., & Heis, A. (2014). Interactive computer-supported learning in mathematics: A comparison of three learning programs on trigonometry. *Journal of Educational Computing Research, 50*(1), 45-65. doi:10.2190/EC.50.1.c

Schaap, L., Schmidt, H. G., & Verkoeijen, P. P. (2012). Assessing knowledge growth in a psychology curriculum: Which students improve most? *Assessment & Evaluation in Higher Education, 37*(7), 875-887. doi:10.1080/02602938.2011.581747

Schaffner, E., & Schiefele, U. (2013). The prediction of reading comprehension by cognitive and motivational factors: Does text accessibility during comprehension testing make a difference? . *Learning and Individual Differences, 26*, 42-54. doi:10.1016/j.lindil".2013.04.003

Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R. (2004). Kindergarten Prediction of Reading Skills: A Longitudinal Comparative Analysis. *Journal of Educational Psychology, 96*(2), 265-282. doi:10.1037/0022-0663.96.2.265

Schmidt, H. K., Rothgangel, M., & Grube, D. (2015). Prior knowledge in recalling arguments in bioethical dilemmas. *Frontiers in Psychology, 6*(1292), 1-10. doi:10.3389/fpsyg.2015.01292

Schneider, W., & Korkel, J. (1989). The knowledge base and text recall: Evidence from a short-term longitudinal study. *Contemporary Educational Psychology, 14*(4), 382-393. doi:10.1016/0361-476X(89)90023-4

Schnotz, W. (1993). Adaptive construction of mental representations in understanding expository texts. *Contemporary Educational Psychology, 18*(1), 114-120. doi:10.1006/ceps.1993.1011

Schnotz, W., & Heis, A. (2009). Semantic scaffolds in hypermedia learning environments. *Computers in Human Behavior, 25*(2), 371-380. doi:10.1016/j.chb.2008.12.016

Schutz, P. A., Drogosz, L. M., White, V. E., & Distefano, C. (1998). Prior knowledge, attitude, and strategy use in an introduction to statistics course. *Learning and Individual Differences, 10*(4), 291-308. doi:10.1016/S1041-6080(99)80124-1

Schwarz, C. V., & White, B. Y. (2005). Metamodeling Knowledge: Developing Students' Understanding of Scientific Modeling. *Cognition and Instruction, 23*(2), 165-205. doi:10.1207/s1532690xci2302_1

Senechal, M., Ouellette, G., & Young, L. (2004). Testing the concurrent and predictive relations among articulation accuracy, speech perception, and phoneme awareness. *Journal of Experimental Child Psychology, 89*(3), 242-269. doi:10.1016/j.jecp.2004.07.005

Shapiro, A. M. (2004). How Including Prior Knowledge As a Subject Variable May Change Outcomes of Learning Research. *American Educational Research Journal, 41*(1), 159-189. doi:10.3102/00028312041001159

Shatil, E., Share, D. L., & Levin, I. (2000). On the contribution of kindergarten writing to grade 1 literacy: A longitudinal study in Hebrew. *Applied Psycholinguistics, 21*(1), 1-21. doi:10.1017/S0142716400001016

Shen, B., & Chen, A. (2006). Examining the interrelations among knowledge, interests, and learning strategies. *Journal of Teaching in Physical Education, 25*(2), 182-199.

Shintani, N. (2017). The effects of the timing of isolated FFI on the explicit knowledge and written accuracy of learners with different prior knowledge of the linguistic target *Studies in Second Language Acquisition, 39*(1), 129-166. doi:10.1017/S0272263116000127

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., . . . Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science, 23*(7), 691-697. doi:10.1177/0956797612440101

Silverman, R. (2013). Investigating video as a means to promote vocabulary for at-risk children. *Contemporary Educational Psychology, 38*(3), 170-179. doi:10.1016/j.cedpsych.2013.03.001

Singh, L., Reznick, S. J., & Liang Xuehua. (2012). Infant word segmentation and childhood vocabulary development: A longitudinal analysis. *Developmental Science, 15*(4), 482-295. doi:10.1111/j.1467-7687.2012.01141.x

Snow, R. E., Tiffin, J., & Seibert, W. F. (1965). Individual differences and instructional film effects. *Journal of Educational Psychology, 56*(6), 315-326. doi:10.1037/h0022788

Sparks, R. L., Patton, J., & Murdoch, A. (2014). Early reading success and its relationship to reading achievement and reading volume: Replication of '10 years later'. *Reading and Writing, 27*(1), 189-211. doi:10.1007/s11145-013-9439-2

Stahl, S. A., Hare, V. C., Sinatra, R., & Gregory, J. F. (1991). Defining the role of prior knowledge and vocabulary in reading comprehension: The retiring of number 4. *Journal of Reading Behavior, 23*(4), 487-508. doi:10.1080/10862969109547755

Stahl, S. A., & Jacobson, M. G. (1986). Vocabulary difficulty, prior knowledge, and text comprehension. *Journal of Reading Behavior, 18*(4), 309-323. doi:10.1080/10862968609547578

Stahl, S. A., Jacobson, M. G., Davis, C. E., & Davis, R. L. (1989). Prior knowledge and difficult vocabulary in the comprehension of unfamiliar text. *Reading Research Quarterly, 24*(1), 27-43. doi:10.2307/748009

Stock, P., Desoete, A., & Roeyers, H. (2007). Early markers for arithmetic difficulties. *Educational and Child Psychology, 24*(2), 28-39.

Stock, P., Desoete, A., & Roeyers, H. (2009). Predicting arithmetic abilities: The role of preparatory arithmetic markers and intelligence. *Journal of Psychoeducational Assessment, 27*(3), 237-251. doi:10.1177/0734282908330587

Stuart, M. (1995). Prediction and qualitative assessment of five- and six-year-old children's reading: A longitudinal study. *British Journal of Educational Psychology, 65*(3), 287-296. doi:10.1111/j.2044-8279.1995.tb01150.x

Taboada, A., & Guthrie, J. T. (2006). Contributions of Student Questioning and Prior Knowledge to Construction of Knowledge From Reading Information Text. *Journal of Literacy Research, 38*(1), 1-35. doi:10.1207/s15548430jlr3801_1

Taft, M. L., & Leslie, L. (1985). The effects of prior knowledge and oral reading accuracy on miscues and comprehension. *Journal of Reading Behavior, 17*(2), 163-179. doi:10.1080/10862968509547537

Tarchi, C. (2010). Reading comprehension of informative texts in secondary school: A focus on direct and indirect effects of reader's prior knowledge. *Learning and Individual Differences, 20*(5), 415-420. doi:10.1016/j.Enctif. 2010.04.002

Tarchi, C. (2017). Comprehending expository texts: the role of cognitive and motivational factors. *Reading Psychology, 38*(2), 154-181. doi:10.1080/02702711.2016.1245229

Thompson, R. A., & Zamboanga, B. L. (2003). Prior knowledge and its relevance to student achievement in introduction to psychology. *Teaching of Psychology, 30*(2), 96-101. doi:10.1207/S15328023TOP3002_02

Thompson, R. A., & Zamboanga, B. L. (2004). Academic Aptitude and Prior Knowledge as Predictors of Student Achievement in Introduction to Psychology. *Journal of Educational Psychology, 96*(4), 778-784. doi:10.1037/0022-0663.96.4.778

Torppa, M., Georgiou, G. K., Lerkkanen, M.-K., Niemi, P., Poikkeus, A.-M., & Nurmi, J.-E. (2016). Examining the simple view of reading in a transparent orthography: A longitudinal study from kindergarten to grade 3. *Merrill-Palmer Quarterly, 62*(2), 179-206. doi:10.13110/merrpalmquar1982.62.2.0179

Tunmer, W. E., & Chapman, J. W. (2012). Does set for variability mediate the influence of vocabulary knowledge on the development of word recognition skills? *Scientific Studies of Reading, 16*(2), 122-140.

Van Blankenstein, F. M., Dolmans, D. H. J. M., Van der Vleuten, C. P. M., & Schmidt, H. G. (2013). Relevant prior knowledge moderates the effect of elaboration during small group discussion on academic achievement. *Instructional Science, 41*(4), 729-744. doi:10.1007/s11251-012-9252-3

van Dijk, A. M., Eysink, T. H. S., & de Jong, T. (2016). Ability-related differences in performance of an inquiry task: The added value of prompts. *Learning and Individual Differences, 47*(2016), 145-155. doi:10.1016/j.lindif.2016.01.008

Van Joolingen, W. R., Aukes, A. V. A., Gijlers, H., & Bollen, L. (2015). Understanding elementary astronomy by making drawing-based models. *Journal of Science Education and Technology, 24*(2-3), 256-264. doi:10.1007/s10956-014-9540-6

Verhagen, W., Aarnoutse, C., & Van Leeuwe, J. (2008). Phonological awareness and naming speed in the prediction of Dutch children's word recognition. *Scientific Studies of Reading, 12*(4), 301-324. doi:10.1080/10888430802132030

Verhagen, W., Aarnoutse, C., & Van Leeuwe, J. (2009). The predictive power of phonemic awareness and naming speed for early Dutch word recognition. *Educational Research and Evaluation, 15*(1), 93-116. doi:10.1080/13803610802663912

Vukovic, R. K., Fuchs, L. S., Geary, D. C., Jordan, N. C., Gersten, R., & Siegler, R. S. (2014). Sources of individual differences in children's understanding of fractions. *Child Development, 85*(4), 1461-1476. doi:10.1111/cdev.12218

Walcott, C. M., Scheemaker, A., & Bielski, K. (2010). A longitudinal investigation of inattention and preliteracy development. *Journal of Attention Disorders, 14*(1), 79-85. doi:10.1177/1087054709333330

Waniek, J., & Schafer, T. (2009). The role of domain and system knowledge on text comprehension and information search in hypermedia. *Journal of Educational Multimedia and Hypermedia, 18*(2), 221-240.

Watts, T. W., Duncan, G. J., Chen, M., Claessens, A., Davis-Kean, P. E., Duckworth, K., . . . Susperreguy, M. I. (2015). The role of mediators in the development of longitudinal mathematics achievement associations. *Child Development, 86*(6), 1892-1907. doi:10.1111/cdev.12416

Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What's past is prologue: Relations between early mathematics knowledge and high school achievement. *Educational Researcher, 43*(7), 352-360. doi:10.3102/0013189X14553660

Webb, S., & Chang, A. C.-S. (2015). How does prior word knowledge affect vocabulary learning progress in an extensive reading program? *Studies in Second Language Acquisition, 37*(04), 1-25 (651-675). doi:10.1017/S0272263114000606

Wecker, C., Rachel, A., Heran-Dorr, E., Waltner, C., Wiesner, H., & Fischer, F. (2013). Presenting theoretical ideas prior to inquiry activities fosters theory-level knowledge. *Journal of Research in Science Teaching, 50*(10), 1180-1206. doi:10.1002/tea.21 106

Weinert, F. E., & Helmke, A. (1998). The neglected role of individual differences in theoretical models of cognitive development. *Learning and Instruction, 8*(4), 309-323. doi:10.1016/S0959-4752(97)00024-8

Wetzels, S. A. J., Kester, L., Van Merrienboer, J. J. G., & Broers, N. J. (2011). The influence of prior knowledge on the retrieval-directed function of note taking in prior knowledge activation. *British Educational Research Journal, 81*(2), 274-291. doi:10.1348/000709910X517425

Wilkins, J. L. M., & Ma, X. (2002). Predicting student growth in mathematical content knowledge. *The Journal of Educational Research, 95*(5), 288-298. doi:10.1080/00220670209596602

Winters, F. I., & Azevedo, R. (2005). High-school students' regulation of learning during computer-based science inquiry. *Journal of Educational Computing Research, 33*(2), 189-217. doi:10.2190/F7HM-9JN5-JUX8-4BM9

Wolfe, M. B. W., & Mienko, J. A. (2007). Learning and memory of factual content from narrative and expository text. *British Journal of Educational Psychology, 77*(3), 541-564. doi:10.1348/000709906X143902

Wolfe, M. B. W., & Woodwyk, J. M. (2010). Processing and memory of information presented in narrative or expository texts. *British Journal of Educational Psychology, 80*(3), 341-362. doi:10.1348/000709910X485700

Wu, X., Lowyck, J., Sercu, L., & Elen, J. (2013). Task complexity, student perceptions of vocabulary learning in EFL, and task performance. *British Journal of Educational Psychology, 83*(1), 160-181. doi:10.1111/j.2044-8279.2011.02061.x

Yeung, P.-s., Ho, C. S.-h., Chan, D. W.-o., & Chung, K. K.-h. (2013). Modeling the relationships between cognitive-linguistic skills and writing in Chinese among elementary grades students. *Reading and Writing, 26*(7), 1195-1221. doi:10.1007/s11145-012-9411-6

Yeung, P.-s., Ho, C. S.-H., Chan, D. W.-o., & Chung, K. K.-h. (2016). Orthographic skills important to Chinese literacy development: the role of radical representation and

orthographic memory of radicals. *Reading and Writing, 29*(9), 1935-1958. doi:10.1007/s11145-016-9659-3

Zacharia, Z. C., Loizou, E., & Papaevripidou, M. (2012). Is physicality an important aspect of learning through science experimentation among kindergarten students? *Early Childhood Research Quartely, 27*(3), 447-457. doi:10. 1O16/j.ccrcsq.20 12.02.004

Zhang, J., McBride-Chang, C., Wagner, R. K., & Chan, S. (2014). Uniqueness and overlap: Characteristics and longitudinal correlates of native Chinese children's writing in English as a foreign language. *Bilingualism: Language and Cognition, 17*(2), 347-363. doi:10.1017/S1366728913000163.

Zhang, X., Koponen, T., Rasanen, P., Aunola, K., Lerkkanen, M.-K., & Nurmi, J.-E. (2014). Linguistic and spatial skills predict early arithmetic development via counting sequence knowledge. *Child Development, 85*(3), 1091-1107. doi:10.1177/0956797613519111

Zhang, Y., Tardif, T., Shu, H., Li, H., Liu, H., McBride-Chang, C., . . . Zhang, Z. (2013). Phonological skills and vocabulary knowledge mediate socioeconomic status effects in predicting reading outcomes for Chinese children. *Developmental Psychology, 49*(4), 665-671. doi:10.1037/a0028612

Zhu, L., & Grabowski, B. L. (2006). Web-Based Animation or Static Graphics: Is the Extra Cost of Animation Worth It? *Journal of Educational Multimedia and Hypermedia, 15*(3), 329-347.

Zumbach, J., Schmitt, S., Reimann, P., & Starkloff, P. (2006). Learning life sciences: Design and development of a virtual molecular biology learning lab. *Journal of Computers in Mathematics and Science Teaching, 25*(3), 281-300.

Appendix D: Coded Moderator Variables (Study 2)

	Posttest knowledge		Absolute knowledge gains		Normalized knowledge gains	
	j	*k*	*j*	*k*	*j*	*k*
Knowledge characteristics						
Knowledge type[a]						
Mixed	74	645	5	16	3	7
Declarative	159	2125	14	40	9	19
Procedural	10	51	2	8	1	6
Knowledge sub-type[a]						
Declarative: facts	41	268	5	11	2	4
Declarative: concepts	128	1244	13	22	8	15
Procedural: cognitive skill	10	51	2	8	1	6
Broad content area[a]						
STEM	87	578	13	40	9	23
Language	99	3055	4	18	2	6
Humanities	5	19	0	0	0	0
Social sciences	17	107	3	4	2	3
Health sciences	2	9	0	0	0	0
Sports	5	54	0	0	1	1
Content domain[a]						
Mathematics	37	416	5	21	3	11
Physics	16	55	2	3	2	3
Chemistry	6	20	0	0	0	0
Biology	21	62	4	10	4	9
Geosciences	6	19	2	6	0	0
Computer sciences	1	1	0	0	0	0
Medicine and nursing	2	9	0	0	0	0
Psychology	14	99	3	4	2	3
First language	94	2715	1	10	0	0
Second language	10	53	3	8	2	6
Sports	5	54	1	1	1	1
History	3	10	0	0	0	0
Other	3	20	0	0	0	0
Similarity of prior knowledge and learning outcome						
Similarity of the knowledge type						

Low	117	1398	0	0	1	1
High	210	2821	20	62	13	32
Similarity of the content domain						
Low	72	688	0	0	0	0
High	209	3535	20	62	14	33
Similarity of the physical context						
Low	7	94	0	0	0	0
High	228	4071	20	62	14	33
Similarity of the temporal context						
Low	184	3974	14	46	10	23
High	55	248	6	16	4	10
Time between assessments	171	3908	16	53	8	21
Similarity of the functional context						
Low	4	12	1	1	1	1
High	230	4204	19	61	13	32
Similarity of the social context						
Low	15	185	0	0	1	1
High	222	3909	19	56	13	32
Similarity of the modality						
Low	46	650	0	0	1	1
High	205	3064	18	55	13	32
Learner characteristics						
Age	136	3362	8	30	5	15
Educational level						
Daycare	5	139	0	0	0	0
Kindergarten/preschool	57	1790	1	10	1	1
Primary education	71	1676	4	15	3	11
Secondary education	49	228	4	10	4	5
Higher education	57	358	7	16	4	11
Continued education	5	22	3	9	2	5
Several	3	10	1	2	0	0
Environmental characteristics						
Country[b]						
Australia	3	29	0	0	0	0
Austria	1	30	0	0	0	0

Belgium	2	80	0	0	0	0
Canada	5	115	0	0	0	0
Peoples Republic of China	5	300	0	0	0	0
Hong Kong	6	168	0	0	0	0
Taiwan	3	9	0	0	0	0
Denmark	2	7	0	0	0	0
Finland	8	322	0	0	0	0
France	2	122	0	0	0	0
Germany	9	253	0	0	0	0
Greece	1	27	0	0	0	0
Israel	4	49	0	0	0	0
Italy	1	36	0	0	0	0
Korea	1	45	0	0	0	0
Netherlands	5	163	1	1	0	0
New Zealand	2	42	0	0	0	0
Norway	3	172	0	0	0	0
Singapore	1	8	0	0	0	0
Spain	2	5	0	0	0	0
Sweden	1	12	0	0	0	0
Turkey	1	2	0	0	0	0
United Kingdom	12	627	0	0	0	0
USA	52	904	1	6	0	0
Intervention Setting						
No intervention	14	150	0	0	0	0
School instruction only	78	2772	2	7	0	0
School instruction and other intervention	57	867	5	20	3	5
Other intervention only	85	397	12	33	11	28
Intervention Duration						
0-2 hours	73	352	10	29	9	26
2-24 hours	7	29	1	4	0	0
2-7 days	5	16	0	0	0	0
> 1 week	57	889	7	24	5	7
Cognitive demands of intervention						
Lower	39	100	8	13	8	12
Higher	38	96	7	11	6	10
Instructional methods in intervention						

Written instruction						
No	23	226	3	9	4	10
Yes	101	463	13	33	9	21
Oral instruction						
No	85	495	11	30	8	20
Yes	36	253	4	12	4	11
Multimedia instruction						
No	70	489	5	19	4	14
Yes	43	149	9	21	7	15
Practice						
No	75	440	7	18	4	10
Yes	47	302	8	23	9	21
Constructive activities						
No	81	501	9	27	7	23
Yes	45	158	7	14	6	7
Technology						
No	78	521	5	19	5	18
Yes	41	127	9	20	6	10
Feedback						
No	94	561	9	24	7	17
Yes	18	65	6	15	5	11
Collaborative learning						
No	98	569	13	38	10	27
Yes	22	87	2	3	3	3
Problem-based learning						
No	96	583	14	39	11	28
Yes	17	93	3	3	2	2
Methodological study characteristics						
Study design						
Group differences in prior knowledge (quasi-experimental or experimental design)	83	403	15	39	13	32
Individual differences in prior knowledge (correlational design)	152	3820	5	23	1	1
Randomized controlled trial						
No	226	4179	20	62	14	33
Yes	9	44	0	0	0	0

Number items T1	186	3053	15	44	12	30
Number items T2	189	3088	14	44	11	29
Response format[c]						
Open	101	1368	6	10	3	5
Fill-in	7	34	2	8	2	8
Single or multiple choice	66	508	7	16	4	9
Rating	1	2	0	0	0	0
Behaviour	1	1	0	0	0	0
Other	6	30	0	0	0	0
Various	18	99	3	9	3	9
Retention test						
No	235	4159	20	62	14	33
Yes	13	64	0	0	0	0
Same test at T1 and T2						
No	201	3606	5	17	2	4
Yes	104	615	17	45	12	29
Measures at T2						
Outcome						
Knowledge	228	4169	19	61	13	32
Achievement	10	54	1	1	1	1
Domain specificity						
Specific domain	232	4179	20	62	14	33
Various domains	8	44	0	0	0	0

[a] Only studies assessing the same specific knowledge characteristic at T1 and T2 are presented.

[b] Only studies including any kind of school instructions are presented.

[c] Only studies using the same response format at T1 and T2 are presented.

Appendix E: Equations of the Statistical Analyses (Study 2)

Computation of effect sizes

Cohen's *d* for posttest knowledge. To calculate Cohen's *d* for group comparisons at T2 we used the following formula (Cumming, 2012):

$$d = \frac{X_{HPK,T2} - X_{LPK,T2}}{\frac{SD^2_{HPK,T2} + SD^2_{LPK,T2}}{2}} \quad (1)$$

in which the numerator is the difference between the means of the high prior knowledge group (HPK) and the low prior knowledge group (LPK) and the denominator is the pooled standard deviation for the means of these groups.

Cohen's *d* for knowledge gains. For the calculation of Cohen's *d* for knowledge gains, we calculated the change of knowledge across time as pretest-adjusted posttest gain difference effect sizes (absolute gain scores) and pretest-adjusted posttest standardized gain difference effect sizes (normalized gain scores). We calculated the pretest-adjusted posttest gain difference effect sizes as follows:

$$d_{gain,abs} = \left[\frac{(X_{HPK,T2} - X_{HPK,T1}) - (X_{LPK,T2} - X_{LPK,T1})}{SD_{pooled,T1,gain}}\right] \quad (2)$$

in which the numerator is the difference in absolute gains for the high prior knowledge group and the low prior knowledge group and the denominator is the pooled standard deviation for the pre-test scores in those groups (using the pooled pretest standard deviation).

We calculated the pretest-adjusted posttest standardized gain difference effect sizes as follows (derived from Morris, 2008, p. 369):

$$d_{gain,re} = \left[\frac{G_{rel,HPK} - G_{rel,LPK}}{SD_{pooled,T1,gain}}\right] \quad (3)$$

in which the numerator is the difference in normalized gains for the high prior knowledge group and the low prior knowledge group and the denominator is the pooled standard deviation for the pre-test scores in those groups (using the pooled pretest standard deviation). The relative difference in gains for each group can be obtained using:

$$G_{rel} = \frac{M_{T2} - M_{T1}}{G_{max}} \quad (4)$$

with

$$G_{max} = M_{max} - M_{T1} \quad (5)$$

in which the nominator is the difference between the obtained scores at the first and second measurement point and the denominator is the maximal possible gain calculated by the difference of the highest possible score of the test and the mean at T1.

The standardized deviations of the normalized gain scores can be calculated as (derived from Morris, 2008, Formula 9):

$$SD_{T1,gain} = \sqrt{\frac{(n_{HPK}-1)\left(\frac{SD_{T1,HPK}}{G_{max,HPK}}\right)^2 + (n_{LPK}-1)\left(\frac{SD_{T1,LPK}}{G_{max,LPK}}\right)^2}{n_{HPK}+n_{LPK}-2}} \quad (6)$$

We transformed all Cohen's *d*s to *r*s using the following formula (Borenstein et al., 2009, p. 48):

$$r = \frac{d}{\sqrt{d^2+a}} \quad (7)$$

with

$$a = \frac{(n_{HPK}+ n_{LPK})^2}{n_{HPK}\times n_{LPK}} \quad (8)$$

Meta-analytic model

To estimate the overall strength of the correlation between prior knowledge and learning outcomes, we estimated a simple RVE meta-regression model:

$$y_{ij} = \beta_0 + u_j + e_{ij} \quad (9)$$

where y_{ij} is the *i*th correlation effect size in the *j*th study, β_0 is the average population effect of the correlation, u_j is the study level random effect such that $\text{Var}(u_j) = \tau^2$ is the between-study variance component, and e_{ij} is the residual for the *i*th effect size in the *j*th study. To estimate the variability in the effect size due to moderator variables, we estimated a mixed-effects RVE meta-regression model:

$$y_{ij} = \beta_0 + \beta_1(Moderator_1)_{ij} + \cdots + \beta_k(Moderator_k)_{ij} + u_j + e_{ij} \quad (10)$$

In this model, each moderator represents a continuous variable or specific dummy coded level of an included moderator variable (e.g., age as a continuous predictor or the two levels CT or longitudinal as predictors pertaining to the moderator study design). Continuous moderator variables were log-transformed to obtain normal distribution. We set the parameter ρ, that is, the assumed inter-correlation of dependent effect sizes to the default value of .80. Sensitivity analyses revealed that results remained the same across all values of ρ (0–1).

To estimate R^2 for the multiple regression models, we calculated the sum across the product of the correlation between the dependent variable with the independent variable and the respective standardized regression coefficient (Tabachnick & Fidell, 2014, p. 167, Fomula 4):

$$R^2 = \sum_{i=1}^{k} r_{yi}\beta_i \quad (11)$$

We calculated the standardized regression coefficients from the unstandardized regression coefficients reported by robumeta with (Bring, 1994, pp. 210, formula 2.3):

$$\beta_i = b_i \times \frac{SD\,(Moderator_i)}{SD\,(r^+)} \quad (12)$$

Appendix F: Mean Effect Sizes and Moderator Analyses Examining Knowledge Gains (Study 2)

	Absolute learning gains								Normalized learning gains							
	j	*k*	r_{AG}^{+}	CI r_{AG}^{+} 95%	τ^2	I^2	Moderator		*j*	*k*	r_{NG}^{+}	CI r_{NG}^{+} 95%	τ^2	I^2	Moderator	
							Sign.	R^2							*Sign.*	R^2
Overall	20	62	-.210	[-.437, .043]	.380	96.99			14	33	-.083	[-.478, .342]	.580	94.83		
Knowledge characteristics																
Similarity of prior knowledge and learning outcome																
Similarity of the temporal context							*ns*	.000							-	-
Low	14	46	-.217	[-.503, .113]	.416	97.83	Ref		10	23	-.170	[-.635, .384]	.381	93.39	-	
High	6	16	-.167	[-.597, .336]	.228	81.11	*ns*		4	10	.220	-	-	-	-	
Time between assessments	16	53					*ns*	.000	8	21					*ns*	.135
Learner characteristics																
Age	8	30					*ns*	.036	5	15					*ns*	.244
Environmental characteristics																
Intervention																
Intervention Duration							*ns*	.016							-	-
0-2 hours	10	29	-.181	[-.489, .168]	.237	86.38	*ns*		9	26	-.069	[-.445, .328]	.326	90.96	-	
> 1 week	7	21	-.359	[-.774, .277]	.405	94.27	Ref		5	7	-.057	-	-	-	-	

Cognitive demands of intervention							*	.281							*	.366
Lower	8	13	-.582	[-.851, -.068]	.502	92.94	Ref		8	12	-.478	[-.804, .064]	.451	93.90	Ref	
Higher	7	11	.106	[-.207, .399]	.086	67.34	*		6	10	.252	[-.130, .569]	.081	81.91	*	
Instructional methods in intervention																
Practice							*ns*	.017							-	-
No	7	18	-.380	[-.786, .250]	.639	93.69	Ref		4	10	-.513	-	-	-	-	
Yes	8	23	-.177	[-.523, .219]	.230	86.08	*ns*		9	21	-.078	[-.481, .352]	.326	90.99	-	
Constructive activities							*ns*	.000							.046	
No	9	27	-.287	[-.633, .155]	.387	87.81	*ns*		7	23	.038	[-.461, .518]	.488	89.99	*ns*	
Yes	7	14	-.297	[-.708, .265]	.413	92.68	Ref		6	7	-.350	[-.789, .324]	.442	93.36	Ref	
Feedback							*ns*	.011							-	-
No	9	24	-.338	[-.686, .135]	.569	92.61	Ref		7	17	-.344	[-.746, .225]	.514	92.65		
Yes	6	15	-.270	[-.778, .449]	.374	88.41	*ns*		5	11	.027	-	-	-	-	
Multivariate effect of all instructional methods	14	39					*ns*	.018								
Methodological study characteristics																
Number of items in prior knowledge measure	15	44					*ns*	.021	12	30					*ns*	.106
Number of items in learning outcome measure	14	44					*ns*	.020	11	29					*ns*	.117
Response format[a]							*ns*	.031							-	-

Open	6	10	-.259	[-.750, .418]	.327	94.44	*ns*		3	5	-.639	-	-	-	-		
Fill-in	2	8	-.281	-	-	-	-		2	8	.446	-	-	-	-		
Single or multiple choice	7	16	-.478	[-.766, -.029]	.247	79.89	Ref		4	9	-.263	-	-	-	-		
Various	3	9	-.048	-	-	-	-		3	9	.310	-	-	-	-		
Same test at T1 and T2							*	.110							-	-	
No	5	17	.208	[.038, .367]	.513	93.35	**		2	4	-.345	-	-	-	-		
Yes	17	45	-.318	[-.549, -.042]	.267	90.77	Ref		12	29	-.029	[-.498, .453]	.749	94.96			

[a] Only for effects in which the moderator had the same level at T1 and T2. For effects with differing moderator levels, see the results category "Similarity of prior knowledge and learning outcome".

[b] Only studies including any kind of school instructions are presented.

Note. * $p < .05$, ** $p < .01$, – df was too small for valid interpretation, we therefore do not report the confidence interval and heterogeneity; ns = nonsignificant. Moderation analysis was calculated for moderating variables with full rows only, respectively.

Appendix G: Latent Transition Probabilities (Study 3)

Latent Transition Probabilities Based on the Estimated Model:

T1 Classes (Rows) by T2 Classes (Columns)

Class					
		C1	C2	C3	C4
C1	Misconceptions profile	.709	.000	.070	.221
C2	Fragmented profile	.000	.389	.611	.000
C3	Indecisive profile	.000	.000	.319	.681
C4	Scientific profile	.000	.000	.000	1.00

Latent Transition Probabilities Based on the Estimated Model:

T2 Classes (Rows) by T3 Classes (Columns)

Class					
		C1	C2	C3	C4
C1	Misconceptions profile	1.00	.000	.000	.000
C2	Fragmented profile	.049	.786	.165	.000
C3	Indecisive profile	.000	.000	.856	.144
C4	Scientific profile	.000	.000	.000	1.00

Latent Transition Probabilities Based On the Estimated Model:

T3 Classes (Rows) by T4 Classes (Columns)

Class		C1	C2	C3	C4
C1	Misconceptions profile	.630	.000	.037	.000
C2	Fragmented profile	.000	.867	.133	.000
C3	Indecisive profile	.000	.039	.765	.196
C4	Scientific profile	.000	.000	.000	1.00

Appendix H: Examples of all three Types of Tasks (Study 3)

Example Task „Chunking“:

Wild horses

Two groups of children participate in the study. The children in one of these groups (Group A) are twelve year-olds, who know little about horses in general. The children in the other group (Group B) are eight year-olds, who know a lot about horses in general. The two groups do not differ in terms of the children’s intelligence and the number of boys and girls in the group. All participants read a simple text about wild horses that teaches them what kinds of species exist, where they are found and what they need to live. Afterwards, the children are asked to answer questions assessing their knowledge and comprehension, based on the text from memory.

How sure are you that each of the following statements is correct or incorrect?

	Definitely incorrect Definitely correct
Group A will perform markedly better than Group B because they have four more years of practice in reading and remembering from texts than the younger children.	o o o o o o o
Group B will perform markedly worse than Group A, because the children are on a lower stage of cognitive development and therefore cannot process information so well.	o o o o o o o
Group B will perform markedly better than Group A because they have more prior knowledge and therefore can store information from the text in a more structured way.	o o o o o o o
Group A will perform markedly worse than Group B because the memory of twelve year-olds is partly impaired already due to hormonal changes during puberty.	o o o o o o o
Both groups will perform almost equally well because they have read the same text with the result that each person has memorized the same information.	o o o o o o o
There barely will be any difference between the two groups because their intelligence is similar and therefore they can process information almost equally well.	o o o o o o o

Example Task „Interference“:

Font Sizes

In one study the participants are presented with 50 pairs of simple numbers successively on a computer screen, for example 4 and 9. One number is always presented in a smaller font size and the other one in a larger font size. The participants' task for each pair of numbers is to answer as quickly as possible whether the left or the right number is presented in a larger font size. For half of the participants (Group A) the number with the larger numeric value (e.g. 9) is always presented in a larger font size and the number with the smaller numeric value (e.g. 4) in a smaller font size. For the other half of the participants (Group B) the number with the smaller numeric value (e.g. 4) is always presented in a larger font size and the number with the larger numeric value (e.g. 9) in a smaller font size. It is analyzed how many milliseconds participants in Group A and Group B need on average to evaluate which one of the two numbers is presented in a larger font size.

How sure are you that each of the following statements that is correct or incorrect?

	Definitely incorrect Definitely correct
Group A will be markedly faster than Group B because the numeric values and font sizes in Group B provide the brain with contradicting information.	o o o o o o o
Group B will be markedly slower than Group A because people are used to larger numeric values being presented in larger font sizes than smaller numeric values in everyday life.	o o o o o o o
Group B will be markedly faster than Group A because it is more distinct when numeric values and font sizes do not match and therefore it is remembered more easily.	o o o o o o o
Group A will be markedly slower than Group B because the task of Group B is more unusual and therefore more interesting and motivating.	o o o o o o o
Both groups will perform almost equally fast because the brain processes numeric values and font sizes independently.	o o o o o o o
There barely will be any difference between the two groups because the task is about font sizes only and the participants will pay no attention to the numbers.	o o o o o o o

Example Task „Source Monitoring“:

Europe

Participants in a study know little about the European economic system. They read a text on the topic “Should all Europeans be allowed to live and work anywhere in the European Union without restrictions?” Participants are told that the author of the text is a participant in a casting show for singing talents and has little expertise in politics. The arguments in the text are partly true and partly incorrect.

The participants rate, right after reading the text, how convincing the arguments are (Test 1). One month later, they are presented with the arguments a second time without announcement and again, they are asked again how convincing they find the arguments (Test 2).

How sure are you that each of the following statements is correct or incorrect?

	Definitely incorrect Definitely correct
The arguments will be judged to be clearly more convincing in Test 1 than in Test 2 because they will be more available in memory during Test 1.	o o o o o o o
The arguments will be judged to be clearly less convincing in Test 2 than in Test 1 because the participants will be less impressed by the author’s fame.	o o o o o o o
The arguments will be judged to be clearly more convincing in Test 2 because participants will have spent more time thinking about the topic and will assume a more positive attitude.	o o o o o o o
The arguments in Test 1 will be judged to be clearly less convincing than in Test 2 because participants will be more likely to remember in Test 1 that the arguments were from an author with little expertise in politics.	o o o o o o o
The arguments are judged in a comparable way in both tests because the arguments did not change in the meantime.	o o o o o o o
The judgment of the arguments will barely differ between the two tests because attitudes towards political topics are embedded deeply in memory.	o o o o o o o

References

Aboelela, S. W., Larson, E., Bakken, S., Carraquillo, O., Formicola, A., Glied, S. A., . . . Gebbie, K. M. (2007). Defining interdisciplinary research: Conclusions from a critical review of the literature. *Health Services Research, 42*(1 Pt 1), 329-346. doi:10.1111/j.1475-6773.2006.00621.x

Abraham, W. C. (2008). Metaplasticity: Tuning synapses and networks for plasticity. *Nature Reviews Neuroscience, 9*, 387. doi:10.1038/nrn2356

Abraham, W. C., & Bear, M. F. (1996). Metaplasticity: The plasticity of synaptic plasticity. *Trends in Neuroscience, 19*(4), 126-130. doi:10.1016/S0166-2236(96)80018-X

Alexander, P. A., & Mayer, R. E. (2011). Introduction to research on instruction. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 245-249). New York, NY: Routledge.

Allison, P. D. (1990). Change scores as dependent variables in regression analysis. *Sociological Methodology, 20*, 93-114. doi:10.2307/271083

Altman, D. G., & Royston, P. (2006). The cost of dichotomising continuous variables. *British Medical Journal, 332*(7549), 1080. doi:10.1136/bmj.332.7549.1080

Amadieu, F., Salmerón, L., Cegarra, J., Paubel, P.-V., Lemarié, J., & Chevalier, A. (2015). Learning from concept mapping and hypertext: An eye tracking study. *Educational Technology & Society, 18*(4), 100-112.

Amadieu, F., Tricot, A., & Marine, C. (2009). Prior knowledge in learning from a non-linear electronic document: Disorientation and coherence of the reading sequences. *Computers in Human Behavior, 25*(2), 381-388. doi:10.1016/j.chb.2008.12.017

American Psychological Association. (2001). *Publication manual* (5th ed.). Washington, DC: APA.

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist, 57*(12), 1060-1073.

Ammann, C., Spampinato, D., & Márquez-Ruiz, J. (2016). Modulating motor learning through transcranial direct-current stimulation: An integrative view. *Frontiers in Psychology, 7*, 1981. doi:10.3389/fpsyg.2016.01981

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*(4), 1036-1060. doi:10.1037/0033-295X.111.4.1036

Andrich, D. (1988). *Rasch models for measurement*. Beverly Hills: Sage Publications.

Ardolino, G., Bossi, B., Barbieri, S., & Priori, A. (2005). Non-synaptic mechanisms underlie the after-effects of cathodal transcutaneous direct current stimulation of the human brain. *The Journal of Physiology, 568*(2), 653-663. doi:10.1113/jphysiol.2005.088310

Arkes, H. R., & Freedman, M. R. (1994). A demonstration of the costs and benefits of expertise in recognition memory. *Memory & Cognition, 12*(1), 84-89. doi:10.3758/BF03197000

Asparouhov, T., & Muthén, B. (2010). Computing the strictly positive Satorra-Bentler chi-square test in Mplus. *Mplus Web Notes, 12*.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89-195). New York: Academic Press.

Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. New York and Toronto: Holt, Rinehart and Winston.

Baddeley, A. D., Eysenck, M. W., & Anderson, M. C. (2009). *Memory*. Hove: Psychology Press.

Badham, S. P., Hay, M., Foxon, N., Kaur, K., & Maylor, E. A. (2016). When does prior knowledge disproportionately benefit older adults' memory? *Aging, Neuropsychology, and Cognition, 23*(3), 338-365. doi:10.1080/13825585.2015.1099607

Bailey, D. H., Duncan, G. J., Watts, T., Clements, D. H., & Sarama, J. (in press). Risky business: Correlation and causation in longitudinal studies of skill development. *American Psychologist*.

Baker, W. L., White, C. M., Cappelleri, J. C., Kluger, J., & Coleman, C. I. (2009). Understanding heterogeneity in meta-analysis: The role of meta-regression. *International Journal of Clinical Practice, 63*(10), 1426-1434. doi:10.1111/j.1742-1241.2009.02168.x

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*(4), 612-637. doi:10.1037//0033-2909.128.4.612

Bartlett, F. C. (1932). Remembering: a study in experimental and social psychology. *British Journal of Educational Psychology, 3*(2), 187-192. doi:10.1111/j.2044-8279.1933.tb02913.x

Bast, J., & Reitsma, P. (1997). Mathew effects in reading: A comparison of latent growth curve models and simplex models with structured means. *Multivariate Behavioral Research, 32*(2), 135-167. doi:10.1207/s15327906mbr3202_3

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., . . . Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*(1), 133-180. doi:10.3102/0002831209345157

Baumert, J., Nagy, G., & Lehmann, R. (2012). Cumulative advantages and the emergence of social and ethnic inequality: Matthew effects in reading and mathematics development within elementary schools? *Child Development, 83*(4), 1347-1367. doi:10.1111/j.1467-8624.2012.01779.x

Bear, M. F., & Malenka, R. C. (1994). Synaptic plasticity: LTP and LTD. *Current Opinion in Neurobiology, 4*(3), 389-399. doi:10.1016/0959-4388(94)90101-5

Beeth, M., Duit, R., Prenzel, M., Ostermeier, C., Tytler, R., & Wickman, P.-O. (2003). Quality development projects in science education. In D. Psillos, P. Kariotoglou, V. Tselfes, E. Hatzikraniotis, G. Fassoulopoulos, & M. Kallery (Eds.), *Science education research in the knowledge-based society* (pp. 447-457): Springer.

Bergman, L. R., & Magnusson, D. (1997). A person-oriented approach in research on developmental psychopathology. *Development and Psychopathology, 9*(2), 291-319. doi:10.1017/S095457949700206X

Biggs, J. (2003). *Teaching for quality learning at university* (2nd ed.). Berkshire, UK: Open University Press.

Bilalić, M., McLeod, P., & Gobet, F. (2008). Inflexibility of experts - reality or myth? Quantifying the Einstellung effect in chess masters. *Cognitive Psychology, 56,* 73-102. doi:10.1016/j.cogpsych.2007.02.001

Bilalić, M., McLeod, P., & Gobet, F. (2010). The mechanism of the Einstellung (set) effect: A pervasive source of cognitive bias. *Current Directions in Psychological Science, 19*(2), 111-115. doi:0.1177/0963721410363571

Bjork, R. A. (1992). Interference and memory. In L. R. Squire (Ed.), *Encyclopedia of learning and memory* (pp. 283-288). New York: Macmillan.

Blakemore, S.-J., & Choudhury, S. (2006). Development of the adolescent brain: Implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry, 47*(3-4), 296-312. doi:10.1111/j.1469-7610.2006.01611.x

Bliss, T. V., & Lømo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *The Journal of Physiology, 232*(2), 331-356. doi:10.1113/jphysiol.1973.sp010273

Blume, B. D., Ford, J. K., Baldwin, T. T., & Huang, J. L. (2010). Transfer of training: A meta-analytic review. *Journal of Management, 36*(4), 1065-1105. doi:10.1177/0149206309352880

Bonate, P. L. (2000). *Analysis of pretest-posttest designs*. Boca Raton: Chapman & Hall.

Boomsma, A. (2000). Reporting analyses of covariance structures. *Structural Equation Modeling, 7*(3), 461-483. doi:10.1207/S15328007SEM0703_6

Borenstein, M. (2009). Effect sizes for continuous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 221-235). New York: Russel Sage Foundation.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothenstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: John Wiley & Sons Ltd.

Boros, K., Poreisz, C., Munchau, A., Paulus, W., & Nitsche, M. A. (2008). Premotor transcranial direct current stimulation (tDCS) affects primary motor excitability in humans. *European Journal of Neuroscience, 27*(5), 1292-1300. doi:10.1111/j.1460-9568.2008.06090.x

Bretz Jr., R. D. (1989). College Grade Point Average as a Predictor of Adult Success: A Meta-Analytic Review and Some Additional Evidence *Public Personnel Management, 18*(1), 11-22. doi: 10.1177/009102608901800102

Bring, J. (1994). How to standardize regression coefficients. *The American Statistician, 48*(3), 209-213.

Brod, G., Lindenberger, U., Wagner, A. D., & Shing, Y. L. (2016). Knowledge acquisition during exam preparation improves memory and modulates memory formation. *Journal of Neuroscience, 36*(31), 8103-8111. doi:10.1523/JNEUROSCI.0045-16.2016

Brod, G., Werkle-Bergner, M., & Shing, Y. L. (2013). The influence of prior knowledge on memory: A developmental cognitive neuroscience perspective. *Frontiers in Behavioral Neuroscience, 7*, Article 139. doi:10.3389/fnbeh.2013.0013

Brooks, L. R. (1967). The suppression of visualization by reading. *Quarterly Journal of Experimental Psychology, 19*(4), 289-299. doi:10.1080/14640746708400105

Buch, E. R., Santarnecchi, E., Antal, A., Born, J., Celnik, P. A., Classen, J., . . . Cohen, L. G. (2017). Effects of tDCS on motor learning and memory formation: A consensus and critical position paper. *Clincial Neurophysiology, 128*(4), 589-603. doi:10.1016/j.clinph.2017.01.004

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Carey, S. (1992). The origin and evolution of everyday concepts. In R. N. Giere (Ed.), *Cognitive models of science* (pp. 89-128). Minneapolis: University of Minnesota Press.

Carey, S., & Spelke, E. S. (1994). Domain-specific knowledge and conceptual change. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 169-200). Cambridge, UK: Cambridge University Press.

Carter, E. C., Kofler, L. M., Forster, D. E., & McCullough, M. E. (2015). A series of meta-analytic tests of the depletion effect: Self-control does not seem to rely on a limited resource. *Journal of Experimental Psychology: General, 144*(4), 796-815. doi:10.1037/xge0000083

Carter, E. C., Schönbrodt, F. D., Gervais, W. M., & Hilgard, J. (2017). Correcting for bias in psychology: A comparison of meta-analytic methods. Retrieved from https://osf.io/rf3ys/

Case, R. (1992). *The mind's staircase: Exploring the conceptual underpinnings of children's thought and knowledge*. Hillsdale: Erlbaum.

Castel, A. D., McCabe, D. P., Roedinger, H. L., & Heitman, J. L. (2007). The dark side of expertise. *Psychological Science, 18*(1), 3-5. doi:10.1111/j.1467-9280.2007.01838.x

Castillo, P. E., Chiu, C. Q., & Carroll, R. C. (2011). Long-term plasticity at inhibitory synapses. *Current Opinion in Neurobiology, 21*(2), 328-338. doi:10.1016/j.conb.2011.01.006

Chang, H.-J., Kang, J., Ham, B.-J., & Lee, Y.-M. (2016). A functional neuroimaging study of the clinical reasoning of medical students. *Advances in Health Sciences Education: Theory and Practice, 21*(5), 969-982. doi:10.1007/s10459-016-9685-6

Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. In G. W. Bower (Ed.), *The Psychology of learning and motivation* (Vol. 16, pp. 1-58). New York: Academic Press.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*(1), 55-81. doi:10.1016/0010-0285(73)90004-2

Cheung, M. W.-L. (2015). *Meta-analysis: A structural equation modeling approach*. Chichester: John Wiley & Sons, Ltd.

Chi, M. T. H. (2006). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The cambridge handbook of expertise and expert performance* (pp. 21-30). Cambridge: Cambridge University Press.

Chi, M. T. H., & Ceci, S. J. (1987). Content knowledge: Its role, representation, and restructuring in memory development. *Advances in Child Development and Behavior, 20*, 91-142. doi:10.1016/S0065-2407(08)60401-2

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 7-75). Hillsdale, NJ: Erlbaum.

Chiu, M. M., Chow, B. W.-Y., & Joh, S. W. (2017). Streaming, tracking and reading achievement: A multilevel analysis of students in 40 countries. *Journal of Educational Psychology, 109*(7), 915-934. doi:10.1037/edu0000188

Chung, H., Lanza, S. T., & Loken, E. (2008). Latent transition analysis: Inference and estimation. *Statistics in Medicine, 27*(11), 1834-1854. doi:10.1002/sim.3130

Chuy, M., Alamargot, D., & Passerault, J.-M. (2012). Interference with prior knowledge while writing from sources: Effects of working memory capacity. *Journal of Cognitive Psychology, 24*(7), 815-828. doi:10.1080/20445911.2012.699035

Clark, S. L., & Muthén, B. (2009). Relating latent class analysis results to variables not included in the analysis. *Submitted for publication.*

Clements, D. H., & Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical Thinking and Learning, 6*, 81-89.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159. doi:10.1037/0033-2909.112.1.155

Cohen Kadosh, R., Levy, N., O'Shea, J., Shea, N., & Savulescu, J. (2012). The neuroethics of non-invasive brain stimulation. *Current Biology, 22*(4), R108-R111. doi:10.1016/j.cub.2012.01.013

Cohen, P. A. (1984). College grades and adult achievement: A research synthesis. *Research in Higher Education, 20*(3), 281-293. doi:10.1007/BF00983503

Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis*. Hoboken, NJ: Wiley.

Cook, R. D., & Weisberg, S. (1982). *Residuals and influence in regression*. New York: Chapman and Hall.

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York: Russel Sage Foundation.

Costanzo, F., Varuzza, C., Rossi, S., Sdoia, S., Varvara, P., Oliveri, M., . . . Menghini, D. (2016). Reading changes in children and adolescents with dyslexia after transcranial direct current stimulation. *NeuroReport, 27*(5), 295-300. doi:10.1097/WNR.0000000000000536

Cox, M., Steegen, A., & De Cock, M. (2016). How aware are teachers of students' misconceptions in astronomy? A qualitative analysis in Belgium. *Science Education International, 27*(2), 277-300.

Craik, F. I. M., & Bialystok, E. (2006). Cognition through the lifespan: Mechanisms of change. *Trends in Cognitive Sciences, 10*(3), 131-138. doi:10.1016/j.tics.2006.01.007

Cribbie, R. A., & Jamieson, J. (2004). Decreases in posttest variance and the measurement of change. *Methods of Psychological Research Online, 9*(1), 37-55.

Cronbach, J., & Furby, L. (1970). How we should measure "change": Or should we? *Psychological Bulletin, 74*(1), 68-80. doi:10.1037/h0029382

Crooks, N. M., & Alibali, M. W. (2013). Noticing relevant problem features: Activating prior knowledge affects problem solving by guiding encoding. *Frontiers in Psychology, 4*, 884. doi:10.3389/fpsyg.2013.00884

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis* New York, NY: Routledge.

Daoudal, G., & Debanne, D. (2003). Long-term plasticity of intrinsic excitability: Learning rules and mechanisms. *Learning and Memory, 10*, 456-465. doi:10.1101/lm.64103

Daoudal, G., Hanada, Y., & Debanne, D. (2002). Bidirectional plasticity of excitatory postsynaptic potential (EPSP)-spike coupling in CA1 hippocampal pyramidal neurons. *Proceedings of the National Academy of Sciences of the United States of America, 99*(22), 14512-14517. doi:10.1073/pnas.222546399

Davis, N. J. (2014). Transcranial stimulation of the developing brain: A plea for extreme caution. *Frontiers in Human Neuroscience, 8*(2014), 600. doi:10.3389/fnhum.2014.00600

Day, S. B., & Goldstone, R. L. (2012). The import of knowledge export: Connecting findings and theories of transfer of learning. *Educational Psychologist, 47*(3), 153-176. doi:10.1080/00461520.2012.696438

de Corte, E. (2010). Historical development in the nature of learning. In H. Dumont, D. Istance, & F. Benavides (Eds.), The nature of learning: using research to inspire practice (pp. 35-68). Paris: OECD Publishing. doi:10.1787/9789264086487-en

de Jong, T., & Ferguson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist, 31*(2), 105-113. doi:10.1207/s15326985ep3102_2

Detterman, D. K., & Sternberg, R. J. (1993). *Transfer on trial: Intelligence, cognition, and instruction*. Westport, CT: Ablex.

DeWitt, M. R., Knight, J. B., & Hicks, J. L. (2012). The effects of prior knowledge on the encoding of episodic contextual details. *Psychonomic Bulletin & Review, 19*(2), 251-257. doi:10.3758/s13423-011-0196-4

Dienes, Z., & Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences, 22*(5), 735-808. doi:10.1017/S0140525X99002186

diSessa, A. A. (2008). A bird's-eye view of the "pieces" vs. "coherence" controversy (from the "pieces" side of the fence). In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 35-60). New York: Routledge.

diSessa, A. A., Gillespie, N. M., & Esterly, J. B. (2004). Coherence versus fragmentation in the development of the concept of force. *Cognitive Science, 28*(6), 843-900. doi:10.1016/j.cogsci.2004.05.003

Dissanayaka, T. D., Zoghi, M., Farrell, M., Egan, G. F., & Jaberzadeh, S. (2017). Sham transcranial electrical stimulation and its effects on corticospinal excitability: A

systematic review and meta-analysis. *Reviews in Neurosciences*, Advance online publication. doi:10.1515/revneuro-2017-0026

Dochy, F., & Alexander, P. A. (1995). Mapping prior knowledge: A framework for discussion among researchers. *European Journal of Psychology of Education, 10*(3), 225-242. doi:10.1007/BF03172918

Dochy, F., Segers, M., & Buehl, M. M. (1999). The relations between assessment practices and outcomes of studies: the case of research on prior knowledge. *Review of Educational Research, 69*(2), 145-186. doi:10.2307/1170673

Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: A meta-analysis. *Learning and Instruction, 13*(5), 533-568. doi:10.1016/S0959-4752(02)00025-7

Dubljevic, V., Saigle, V., & Racine, E. (2014). The rising tide of tDCS in the media and academic literature. *Neuron, 82*(4), 731-736. doi:10.1016/j.neuron.2014.05.003

Duff, D., Tomblin, J. B., & Catts, H. (2015). The influence of reading on vocabulary growth: A case for a Matthew effect. *Journal of Speech, Language, and Hearing Research, 58*, 853-864. doi:10.1044/2015_JSLHR-L-13-0310

Duit, R., Treagust, D. F., & Widodo, A. (2008). Teaching science for conceptual change: Theory and practice. *International handbook of research on conceptual change*, 629-646.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., . . . Duckworth, K. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428-1446. doi:0.1037/0012-1649.43.6.1428

Duncan, T. E., Duncan, S. C., Strycker, L. A., Li, F., & Alpert, A. (1999). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications*. Mahwah, NJ: Lawrence Erlbaum Associates.

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-Analysis of experiments with matched groups or repeated measures designs. *Psychological Methods, 1*(2), 170-177. doi:10.1037/1082-989X.1.2.170

Duval, S., & Tweedie, R. (2000). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95*(449), 89-98. doi:10.2307/2669529

Edelsbrunner, P. A., Schalk, L., Schumacher, R., & Stern, E. (2015). *Pathways of conceptual change: Investigating the influence of experimentation skills on conceptual knowledge development in early science education.* Paper presented at the Proceedings of the 37th Annual Conference of the Cognitive Science Society, 620.

Edelsbrunner, P. A., Schalk, L., Schumacher, R., & Stern, E. (under review). Variable control and conceptual change: A large-scale quantitative study in elementary school. *Learning and Individual Differences*. doi:10.17605/OSF.IO/ESCQA

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*(7109), 629-634. doi:10.1136/bmj.315.7109.629

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist, 49*(8), 725-747.

Ericsson, K. A., Chase, W. G., & Faloon, S. (1980). Acquisition of a memory skill. *Science, 208*(4448), 1181-1182. doi:10.1126/science.7375930

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363-406. doi:10.1037/0033-295X.100.3.363

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology, 47*, 273-305. doi:10.1146/annurev.psych.47.1.273

European Commission. (2008). *The European Qualifications Framework for Lifelong Learning (EQF)*. Luxembourg: Office for Official Publications of the European Communities.

Fan, X., & Chen, M. (2001). Parental involvement and students' academic achievement: A meta-analysis. *Educational Psychology Review, 13*(1), 1-22. doi:1040-726X/01/0300-0001

Farnham-Diggory, S. (1994). Paradigms of knowledge and instruction. *Review of Educational Research, 64*(3), 463-477. doi:10.3102/00346543064003463

Ferguson, C. J., & Brannick, M. T. (2011). Publication bias in psychological science: Prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. *Psychological Methods, 17*(1), 120-128. doi:10.1037/a0024445

Fisher, F., Kollar, I., Ufer, S., Sodian, B., Hussmann, H., Pekrun, R., . . . Eberle, J. (2014). Scientific reasoning and argumentation: Advancing an interdisciplinary research agenda in education. *Frontline Learning Research, 2*(3), 28-45. doi:10.14786/flr.v2i3.96

Fisher, M., & Keil, F. C. (2016). The curse of expertise: When more knowledge leads to miscalibrated explanatory insight. *Cognitive Science, 40*(5), 1251-1269. doi:10.1111/cogs.12280

Fisher, Z., & Tipton, E. (2014). Robumeta: Robust variance meta-regression. Retrieved from http://cran.rproject.org/web/packages/robumeta/index.html

Flaig, M., Simonsmeier, B. A., Mayer, A.-K., Rosman, T., Gorges, J., & Schneider, M. (in press). Conceptual change and knowledge integration as learning processes in higher education: A latent transition analysis. *Learning and Individual Differences*.

Flöel, A., Rösser, N., Michka, O., Knecht, S., & Breitenstein, C. (2008). Noninvasive Brain Stimulation Improves Language Learning. *Journal of Cognitive Neuroscience, 20*(8), 1415–1422. doi:10.1162/jocn.2008.20098

Floyer-Lea, A., Wylezinska, M., Kincses, T., & Matthews, P. M. (2006). Rapid modulation of GABA concentration in human sensorimotor cortex during motor learning. *Journal of Neurophysiology, 95*(3), 1639-1644. doi:10.1152/jn.00346.2005

Friege, G., & Lind, G. (2006). Types and qualities of knowledge and their relations to problem solving in physics. *International Journal of Science and Mathematics Education, 4*(3), 437-465. doi:10.1007/s10763-005-9013-8

Fritsch, B., Reis, J., Martinowich, K., Schambra, H. M., Ji, Y., Cohen, L. G., & Lu, B. (2011). Direct current stimulation promotes BDNF-dependent synaptic plasticity: Potential implications for motor learning. *Neuron, 66*(2), 198-204. doi:10.1016/j.neuron.2010.03.035

Fyfe, E. R., & Rittle-Johnson, B. (2016). Feedback both helps and hinders learning: The causal role of prior knowledge. *Journal of Educational Psychology, 108*(1), 82-97. doi:10.1037/edu0000053

Fyfe, E. R., Rittle-Johnson, B., & DeCaro, M. S. (2012). The effects of feedback during exploratory mathematics problem solving: Prior knowledge matters. *Journal of Educational Psychology, 104*(4), 1094-1108. doi:10.1037/a0028389

Gagné, F. (2004). Transforming gifts into talents: The DMGT as a developmental theory. *High Ability Studies, 15*(2), 119-147. doi:10.1080/1359813042000314682

Gaiarsa, J.-L., Caillard, O., & Ben-Ari, Y. (2002). Long-term plasticity at GABAergic and glycinergic synapses: Mechanisms and functional significance. *Trends in Neuroscience, 25*(11), 564-570. doi:10.1016/S0166-2236(02)02269-5

Garry, M., Loftus, E. F., & Brown, S. W. (1994). Memory: A river runs through it. *Consciousness and Cognition, 3*(3-4), 438-451. doi:10.1006/ccog.1994.1025

Gentner, D., Brem, S., Ferguson, R. W., Markman, A. B., Levidow, B. B., Wolff, P., & Forbus, K. D. (1997). Analogical reasoning and conceptual change: A case study of

Johannes Kepler. *Journal of the Learning Sciences, 6*(1), 3-40. doi:10.1207/s15327809jls0601_2

Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction, 16*(6), 511-525. doi:10.1016/j.learninstruc.2006.10.001

Giofrè, D., Borella, E., & Mammarella, I. C. (2017). The relationship between intelligence, working memory, academic self-esteem, and academic achievement. *Journal of Cognitive Psychology, 29*(6), 731-747. doi:10.1080/20445911.2017.1310110

Gobet, F. (1998). Expert memory: A comparison of four theories. *Cognition, 66*(2), 115-152. doi:10.1016/S0010-0277(98)00020-1

Gobet, F., Lane, P. C. R., Croker, S., Cheng, P. C. H., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences, 5*(6), 236-243. doi:10.1016/S1364-6613(00)01662-4

Goff, L. M., & Roediger, H. L. (1998). Imagination inflation for action events: Repeated imaginings lead to illusory recollections. *Memory & Cognition, 26*(1), 20-33.

Goldstone, R. L., & Kersten, A. (2003). Concepts and categorization. In I. B. Weiner (Ed.), *Handbook of psychology* (Vol. 4, pp. 599-621). Hoboken, NJ: Wiley.

Goldwater, M. B., & Schalk, L. (2016). Relational categories as a bridge between cognitive and educational research. *Psychological Bulletin, 142*(7), 729-757. doi:10.1037/bul0000043

Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms and the theory theory. *Psychological Bulletin, 138*(6), 1085-1108. doi:10.1037/a0028044

Greenhalgh, T. (1997). Papers that summarise other papers (systematic reviews and meta-analyses). *British Medical Journal, 315*(7109), 672-675. doi:10.1136/bmj.315.7109.672

Guzzetti, B., Snyder, T. E., Glass, G. V., & Gamas, W. S. (1993). Promoting conceptual change in science: A comparative meta-analysis of instructional interventions from reading education and science education. *Reading Research Quarterly, 28*(2), 116-159. doi:10.2307/747886

Hagger, M. S., N. L. D. Chatzisarantis, Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R., . . . Zwienenberg, M. (2016). A multilab preregistered replication of the ego-depletion effect. *Perspectives on Psychological Science, 11*(4), 546-573. doi:10.1177/1745691616652873

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66*(1), 64-74. doi:10.1119/1.18809

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences, 21*(6), 803-831. doi:10.1017/S0140525X98001769

Hall, C. C., Ariss, L., & Todorov, A. (2007). The illusion of knowledge: When more information reduces accuracy and increases confidence. *Organizational Behavior and Human Decision Making, 103*(2), 227-290. doi:10.1016/j.obhdp.2007.01.003

Hambrick, D. Z., & Engle, R. W. (2002). Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. *Cognitive Psychology, 44*(4), 339-387. doi:10.1006/cogp.2001.0769

Hardy, I., Jonen, A., Möller, K., & Stern, E. (2006). Effects of instructional support within constructivist learning environments for elementary school students' understanding of "floating and sinking". *Journal of Educational Psychology, 98*(2), 307-326.

Hattie, J. A. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London, UK: Routledge.

Hattie, J. A. (2015). The applicability of visible learning to higher education. *Scholarship of teaching and Learning in Psychology, 1*(1), 79-91. doi:10.1037/stl0000021

Hawking, S. (Writer) & S. Hawking (Director). (2013). Designer Human [Video]. In Handel Producations (Producer), *Stephen Hawking's Brave New World Season 2 Episode 5*. Canada: Discovery Channel Canada.

Hecht, H., & Proffitt, D. R. (1995). The price of expertise: Effects of experience on the water-level task. *Psychological Science, 6*(2), 90-95. doi:10.1111/j.1467-9280.1995.tb00312.x

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis methods, 1*(2), 39-65. doi:10.1002/jrsm.5

Hemmer, P., & Persaud, K. (2014). Interaction between categorical knowledge and episodic memory across domains. *Frontiers in Psychology, 5*, 584. doi:10.3389/fpsyg.2014.00584

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relations between physical and semantic size in comparison tasks. *Memory & Cognition, 10*(4), 389-395. doi:10.3758/BF03202431

Henmi, M., & Copas, J. B. (2010). Confidence intervals for random effects meta-analysis and robustness to publication bias. *Statistics in Medicine, 29*(29), 2969-2983. doi:0.1002/sim.4029

Hewson, P. W., Tabachnick, B. R., Zeichner, K. M., & Lemberger, J. (1999). Educating prospective teachers of biology: Findings, limitations, and recommendations. *Science Education, 83*(3), 373-384.

Hickendorff, M., Edelsbrunner, P. A., McMullen, J., Schneider, M., & Trezise, K. (in press). Informative tools for characterizing individual differences in learning: Latent class,

latent profile, and latent transition analysis. *Learning and Individual Differences*. doi:10.17605/OSF.IO/JMTCV

Hiebert, J., Stigler, J. W., & Manaster, A. B. (1999). Mathematical features of lessons in the TIMSS video study. *ZDM – International Journal on Mathematics Education, 31*(6), 196-201. doi:10.1007/BF02652695

Higgins, J. P. T., & Green, S. (2008). *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley-Blackwell.

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*(11), 1539-1558. doi:10.1002/sim.1186

Higgins, S. (2016). Meta-synthesis and comparative meta-analysis of education research findings: Some risks and benefits. *Review of Education, 4*(1), 31-53. doi:10.1002/rev3.3067

Hill, A. T., Fitzgerald, P. B., & Hoy, K. E. (2016). Effects of anodal transcranial direct current stimulation on working memory: A systematic review and meta-analysis of findings from healthy and neuropsychiatric populations. *Brain Stimulation, 9*(2), 197-208. doi:10.1016/j.brs.2015.10.006

Hindman, A. H., Erhart, A. C., & Wasik, B. A. (2002). Reducing the Matthew effect: Lessons from the ExCELL head start intervention. *Early Education and Development, 23*(5), 781-806. doi:10.1080/10409289.2010.549443

Hirschfeld, L. A., & Gelman, S. A. (Eds.). (1994). *Mapping the mind: Domain specificity in cognition and culture*. New York: Cambridge University Press.

Hoffert, S. P. (1997). Meta-analysis is gaining status in science and policymaking. *The Scientist, 11*(18), 1-6.

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

Horvath, J. C., Forte, J. D., & Carter, O. (2015a). Evidence that transcranial direct current stimulation (tDCS) generates little-to-no reliable neurophysiologic effect beyond MEP amplitude modulation in healthy human subjects: A systematic review. *Neuropsychologia, 66*, 213-236. doi:10.1016/j.neuropsychologia.2014.11.021

Horvath, J. C., Forte, J. D., & Carter, O. (2015b). Quantitative review finds no evidence of cognitive effects in healthy populations from single-session transcranial direct current stimulation (tDCS). *Brain Stimulation, 8*(3), 535-550. doi:10.1016/j.brs.2015.01.400

Hovland, C. I., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. *Public opinion quarterly, 15*(4), 635-650. doi:10.1086/266350

Hoyle, R. H., & Isherwood, J. C. (2013). Reporting results from structural equation modeling analyses in Archives of Scientific Psychology. *Archives of Scientific Psychology, 1*, 14-22. doi:10.1037/arc0000004

Huang, F. L., Moon, T. R., & Boren, R. (2014). Are the reading rich getting richer? Testing for the presence of the Matthew effect. *Reading & Writing Quarterly, 30*(2), 95-115. doi:10.1080/10573569.2013.789784

Hulbert, J. C., & Anderson, M. C. (2008). The role of inhibition in learning. In A. S. Benjamin, J. S. De Belle, B. Etnyre, & T. A. Polk (Eds.), *Human learning: biology, brain, and neuroscience*. Oxford: Elsevier.

Hunter, J. E., & Schmidt, F. L. (1996). Cumulative research knowledge and social policy formulation: The critical role of meta-analysis. *Psychology, Public Policy, and Law, 2*(2), 324-347. doi:10.1037/1076-8971.2.2.324

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: Sage.

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. doi:10.1177/0018720812469985

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*(2), 139-155. doi:10.1037/0033-2909.107.2.139

Irwin, J. R., & McClelland, G. H. (2003). Negative consequences of dichotomizing continuous predictor variables. *Journal of Marketing Research, 40*(3), 366-371. doi:10.1509/jmkr.40.3.366.19237

Iuculano, T., & Cohen Kadosh, R. (2013). The mental cost of cognitive enhancement. *Journal of Neuroscience, 33*(10), 4482-4486. doi:10.1523/JNEUROSCI.4927-12.2013

Iuculano, T., & Cohen Kadosh, R. (2014). Preliminary evidence for performance enhancement following parietal lobe stimulation in developmental dyscalculia. *Frontiers in Human Neuroscience, 8,* Article 38. doi:10.3389/fnhum.2014.00038

Jacobson, L., Koslowsky, M., & Lavidor, M. (2012). tDCS polarity effects in motor and cognitive domains: A meta-analytical review. *Experimental Brain Research, 216*(1), 1-10. doi:10.1007/s00221-011-2891-9

Jaekel, J., Eryigit-Madzwamuse, S., & Wolke, D. (2016). Preterm toddlers' inhibitory control abilities predict attention regulation and academic achievement at age 8 years. *The Journal of Pediatrics, 169,* 87-92.e81. doi:10.1016/j.jpeds.2015.10.029

Jensen, C., & Hansen, C. (1995). The effect of prior knowledge on EAP listening-test performance. *Language Testing, 12*(1), 99-119. doi:10.1177/026553229501200106

Johnson, A. (2003). Procedural memory and skill acquisition. In I. B. Weiner (Ed.), *Handbook of Psychology* (pp. 499-523). Hoboken, NJ: Wiley.

Johnson, M. A., & Lawson, A. E. (1998). What are the relative effects of reasoning ability and prior knowledge on Biology achievement in expository and inquiry classes? *Journal of Research in Science Teaching, 35,* 89-103.

Johnson, M. K., Hashtroudi, S., & Lindsay, S. D. (1993). Source monitoring. *Psychological Bulletin, 114*(1), 3-28.

Jonckheere, A. R. (1954). A distribution-free k-sample test against ordered alternatives. *Biometrika, 41*(1/2), 133-145.

Kainulainen, M., McMullen, J., & Lehtinen, E. (2017). Early developmental trajectories toward concepts of rational numbers. *Cognition and Instruction, 35*(1). doi:10.1080/07370008.2016.1251287

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review, 19*(4), 509-539. doi:10.1007/s10648-007-9054-3

Kalyuga, S. (2008). Relative effectiveness of animated and static diagrams: An effect of learner prior knowledge. *Computers in Human Behavior, 24*(3), 852-861. doi:10.1016/j.chb.2007.02.018

Kalyuga, S. (2013). Effects of learner prior knowledge and working memory limitations on multimedia learning. *Procedia - Social and Behavioral Sciences, 83*(4), 25-29. doi:/10.1016/j.sbspro.2013.06.005

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*(1), 23-31. doi:10.1207/S15326985EP3801_4

Keil, F. C. (1996). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.

Kempe, C., Eriksson-Gustavsson, A. L., & Samuelsson, S. (2011). Are there any Matthew effects in literacy and cognitive development? *Scandinavian Journal of Educational Research, 55*(2), 181-196. doi:10.1080/00313831.2011.554699

Kepinska, O., de Rover, M., Caspers, J., & Schiller, N. O. (2017). On neural correlates of individual differences in novel grammar learning: An fMRI study. *Neuropsychologia, 98*, 156-168. doi:10.1016/j.neuropsychologia.2016.06.014

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist, 49*(4), 294-303.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75-86. doi:10.1207/s15326985ep4102_1

Kline, P. (1986). *A handbook of test construction: Introduction to psychometric design*. London: Methuen.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 154-284. doi:10.1037/0033-2909.119.2.254

Krause, B., & Cohen Kadosh, R. (2013). Can transcranial electrical stimulation improve learning difficulties in atypical brain development? A future possibility for cognitive training. *Developmental Cognitive Neuroscience, 6*(100), 176-194. doi:10.1016/j.dcn.2013.04.001

Krause, U.-M., Stark, R., & Mandl, H. (2009). The effects of cooperative learning and feedback on e-learning in statistics. *Learning and Instruction, 19*(2), 158-170. doi:10.1016/j.learninstruc.2008.03.003

Krems, J. (1994). *Wissensbasierte Urteilsbildung*. Bern: Huber.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the graduate record examniation: Implications for graduate student selection and performance. *Psychological Bulletin, 127*(1), 162-181. doi:10.1037/0033-2909.127.1.162

Kyun, S. A., & Lee, H. (2009). The effects of worked examples in computer-based instruction: Focus on the presentation format of worked examples and prior knowledge of learners. *Asia Pacific Education Review, 10*(4), 495-503. doi:10.1007/s12564-009-9044-x

La Paro, K. M., & Pianta, R. C. (2000). Predicting children's competence in the early school years: a meta-analytic review *Review of Educational Research, 70*(4), 443-484.

Laird, J. E. (2012). *The SOAR cognitive architecture*. Cambridge, MA: MIT Press.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*, Article 963. doi:10.3389/fpsyg.2013.00863

Lanza, S. T., & Collins, L. M. (2008). A new SAS procedure for latent transition analysis: Transitions in dating and sexual risk behavior. *Developmental Psychology, 44*(2), 446-456. doi:10.1037/0012-1649.44.2.446

Lanza, S. T., Patrick, M. E., & Maggs, J. L. (2010). Latent transition analysis: Benefits of a latent variable approach to modeling transitions in substance use. *Journal of Drug Issues, 40*(1), 93-120.

Ledford, H. (2015). How to solve the world's biggest problems. *Nature, 525*(7569), 308-311. doi:10.1038/525308a

Lejeune, C., Desmottes, L., Catale, C., & Meulemans, T. (2015). Age difference in dual-task interference effects on procedural learning in children. *Journal of Experimental Child Psychology, 129*(January 2015), 165-172. doi:10.1016/j.jecp.2014.07.007

Lewandowsky, S., & Kirsner, K. (2000). Knowledge partitioning: Context-dependent use of expertise. *Memory & Cognition, 28*(2), 295-305. doi:10.3758/BF03213807

Liebetanz, D., Nitsche, M. A., Tergau, F., & Paulus, W. (2002). Pharmacological approach to the mechanisms of transcranial DC-stimulation–induced after-effects of human motor cortex excitability. *Brain, 125*(Pt 10), 2238-2247. doi:10.1093/brain/awf238

Linn, M. C. (2006). The knowledge integration perspective on learning and instruction. In R. K. Sawyer (Ed.), *The Cambridge Handbook of: The Learning Sciences* (Vol. xix). New York: Cambridge University Press.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Liu, Z. h.-X., Grady, C., & Moscovitch, M. (2017). Effects of prior-knowledge on brain activation and connectivity during associative memory encoding. *Cerebral Cortex, 27*(3), 1991-2009. doi:10.1093/cercor/bhw047

Lix, L. M., & Keselman, H. (1998). To trim or not to trim: Tests of location equality under heteroscedasticity and nonnormality. *Educational and Psychological Measurement, 58*(3), 409-429. doi:10.1177/0013164498058003004

Llabre, M. M., Spitzer, S. S., Saab, P. G., Ironson, G. H., & Schneiderman, N. (1991). The reliability and specificity of delta versus residualized change as measures of cardiovascular reactivity to behavioral challenges. *Psychophysiology, 28*(6), 701-711. doi:10.1111/j.1469-8986.1991.tb01017.x

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology, 7*(4), 560-572.

Lombardi, D., Sinatra, G. M., & Nussbaum, E. M. (2013). Plausibility reappraisals and shifts in middle school students' climate change conceptions. *Learning and Instruction, 27,* 50-62. doi:10.1016/j.learninstruc.2013.03.001

Lonigan, C. J., Burgees, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental Psychology, 36*(5), 596-613. doi:10.1037/0012-1649.36.5.596

Looi, C. Y., Lim, J., Sella, F., Lolliot, S., Duta, M., Avramenko, A. A., & Cohen Kadosh, R. (2017). Transcranial random noise stimulation and cognitive training to improve learning and cognition of the atypically developing brain: A pilot study. *Scientific Reports, 7*(1). doi:10.1038/s41598-017-04649-x

Loyens, S. M. M., Jones, S. H., Mikkers, J., & van Gog, T. (2015). Problem-based learning as a facilitator of conceptual change. *Learning and Instruction, 38,* 34-42. doi:10.1016/j.learninstruc.2015.03.002

Luchins, A. S., & Luchins, E. H. (1959). *Rigidity of behavior: A variational approach to the effect of Einstellung*. Eugene, OR: University of Oregon Press.

Luchins, A. S., & Luchins, E. H. (1987). Einstellung effects. *Science, 238*(4827), 598.

Lynn, S. J., & McConkey, K. M. (Eds.). (1998). *Truth in memory*. New York: Guilford.

Machery, E. (2010). Précis of doing without concepts. *Behavioral and Brain Sciences, 33*(2-3), 195-244. doi:10.1017/S0140525X09991531

Maguire, E. A., Frith, C. D., & Morris, R. G. M. (1999). The functional neuroanatomy of comprehension and memory: The importance of prior knowledge. *Brain: A Journal of Neurology, 122,* 1839-1850. doi:10.1093/brain/122.10.1839

Mancuso, L. E., Ilieva, I. P., Hamilton, R. H., & Farah, M. J. (2016). Does transcranial Direct Current Stimulation improve healthy working memory? A meta-analytic review. *Journal of Cognitive Neuroscience, 28*(8), 1063-1089. doi:10.1162/jocn_a_00956

Maris, E. (1998). Covariance adjustment versus gain scores - revisited. *Psychological Methods, 3,* 309-327. doi:10.1037/1082-989X.3.3.309

Markman, A. B., & Dietrich, E. (2000). Extending the classical view of representation. *Trends in Cognitive Sciences, 4*(12), 470-475. doi:10.1016/S1364-6613(00)01559-X

Masson, S., Potvin, P., Riopel, M., & Brault-Foisy, L.-M. (2014). Differences in brain activation between novices and experts in science during a task involving a common misconception in electricity. *Mind, Brain, and Education, 8*(1), 44-55. doi:10.1111/mbe.12043

May, K., & Hittner, J. B. (2010). Reliability and validity of gain scores considered graphically. *Perceptual and Motor Skills, 111,* 399-406. doi:10.2466/03.PMS.111.5.399-406

Mayer, R. E. (2001). *Multimedia learning*. Cambridge: Cambridge University Press.

Mayer, R. E., Bove, W., Bryman, A., Mars, R., & Tapangco, L. (1996). When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons. *Journal of Educational Psychology, 88*(1), 64-73. doi:10.1037/0022-0663.88.1.64

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 43-52. doi:10.1207/S15326985EP3801_6

McCuin, J. L., Hayhoe, K., & Hayhoe, D. (2014). Comparing the effects of traditional vs. misconceptions-based instruction on student understanding of the greenhouse effect. *Journal of Geoscience Education, 62*(3), 445-459. doi:10.5408/13-068.1

McMullen, J., Laakkonen, E., Hannula-Sormunen, M., & Lehtinen, E. (2015). Modeling the developmental trajectories of rational number concept (s). *Learning and Instruction, 37*, 14-20. doi:10.1016/j.learninstruc.2013.12.004

McMullen, J., Van Hoof, J., Degrande, T., Verschaffel, L., & Van Dooren, W. (under review). Profiles of rational number knowledge in Finnish and Flemish students – A multigroup latent class analysis. *Learning and Individual Differences*.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100*(2), 254-278. doi:10.1037/0033-295X.100.2.254

Medin, D. L., & Ross, B. H. (1992). *Cognitive Psychology*. Fort Worth, TX: Harcourt, Brace Jovanovich College Publishers.

Melby-Lervåg, M., & Lervåg, A. (2011). Cross-linguistic transfer of oral language, decoding, phonological awareness and reading comprehension: A meta-analysis of the correlational evidence. *Journal of Research in Reading*, *34*(1), 114-135. doi:10.1111/j.1467-9817.2010.01477.x

Melcher, J. M., & Schooler, J. W. (1996). The misremembrance of wines past: Verbal and perceptual expertise differentially mediate verbal overshadowing of taste memory. *Journal of memory and language*, *35*(2), 231-245. doi:10.1006/jmla.1996.0013

Merz, C. J., Dietsch, F., & Schneider, M. (2016). The impact of psychosocial stress on conceptual knowledge retrieval. *Neurobiology of Learning and Memory, 134*, 392-399. doi:10.1016/j.nlm.2016.08.020

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81-97. doi:10.1037/0033-295X.101.2.343

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology, 110*(1), 40-48. doi:10.1037//0021-843X.110.1.40

Minarik, T., Berger, B., Althaus, L., Bader, V., Biebl, B., Brotzeller, F., . . . Sauseng, P. (2016). The Importance of Sample Size for Reproducibility of tDCS Effects. *Frontiers in Human Neuroscience, 10*, 453. doi:10.3389/fnhum.2016.00453

Minichiello, L. (2009). TrkB signalling pathways in LTP and learning. *Nature Reviews Neuroscience, 10*(12), 850-860. doi: 10.1038/nrn2738

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med, 6*(7), e1000097. doi:10.1371/journal.pmed.1000097

Moreno-Duarte, I., Gebodh, N., Schestatsky, P., Guleyupoglu, B., Reato, D., Bikson, M., & Fregni, F. (2014). Transcranial electrical stimulation: Transcranial direct current stimulation (tDCS), transcranial alternating current stimulation (tACS), transcranial pulsed current stimulation (tPCS), and transcranial random noise stimulation (tRNS). In R. Cohen Kadosh (Ed.), *The stimulated brain: cognitive enhancement using non-invasive brain stimulation*. London: Academic Press.

Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods, 11*(2), 364-386. doi:10.1177/1094428106291059

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7*(1), 105-125. doi:10.1037//1082-989X.7.1.105

Morton, S. C., Adams, J. L., Suttorp, M. J., & Shekelle, P. G. (2004). *Meta-regression approaches: What, why, when, and how?* Rockville, MD: Agency for Healthcare Research and Quality.

Mua, J.-S., Liac, W.-P., Yaob, Z.-B., & Zhuoc, X.-F. (1999). Deprivation of endogenous brain-derived neurotrophic factor results in impairment of spatial learning and memory in adult rats. *Brain Research, 835*(2), 259-265. doi:10.1016/S0006-8993(99)01592-9

Müller, B. (1999). Use specifity of cognitive skills: Evidence for production rules? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(1), 191-207. doi:10.1037/0278-7393.25.1.191

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, Mass: MIT Press.

Muthén, L. K., & Muthén, B. O. (1998-2015). *Mplus User's Guide. Seventh Edition*. Los Angeles, CA: Muthén & Muthén.

Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research, 76*(3), 413-448. doi:10.3102/00346543076003413

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Newman, L. I., & Polk, T. A. (2008). The computational cognitive neuroscience of learning and memory: Principles and models. In M. Guadagnoli (Ed.), *Human learning: Biology, brain and neuroscience* (Vol. 139, pp. 77-100). Oxford, UK: Elsevier Science.

Niedźwieńska, A., Neckar, J., & Baran, B. (2007). Development and validation of the implicit memory theory scale. *European Journal of Psychological Assessment, 23*(3), 185-192.

Nitsche, M. A., Fricke, K., Henschke, U., Schlitterlau, A., Liebetanz, D., Lang, N., . . . Paulus, W. (2003). Pharmacological modulation of cortical excitability shifts induced by transcranial direct current stimulation in humans. *Journal of Physiology, 533*(1), 293-301. doi:10.1113/jphysiol.2003.049916

Nitsche, M. A., & Paulus, W. (2000). Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation. *Journal of Physiology, 527*(3), 633-639. doi: 10.1111/j.1469-7793.2000.t01-1-00633.x

Nitsche, M. A., Seeber, A., Frommann, K., Klein, C. C., Rochford, C., Nitsche, M. S., . . . Tergau, F. (2005). Modulating parameters of excitability during and after transcranial direct current stimulation of the human motor cortex. *The Journal of Physiology, 568*(Pt 1), 291-303. doi:10.1113/jphysiol.2005.092429

Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 510-520. doi:10.1037/0278-7393.14.3.510

Nurmi, J.-E. (2012). Students' characteristics and teacher–child relationships in instruction: A meta-analysis. *Educational Research Review, 7*(3), 177-197. doi:10.1016/j.edurev.2012.03.001

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

O'Donnell, A., M., Dansereau, D. F., & Hall, R. H. (2002). Knowledge maps as scaffolds for cognitive processing. *Educational Psychology Review, 14*(1), 71-86. doi:10.1023/A:1013132527007

O'Reilly, T., & McNamara, D. S. (2007). The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional "high-stakes" measures of high school students' science achievement. *American Educational Research Journal, 44*, 161-196.

OECD. (2001). *Knowledge and skills for life. First results from the OECD programme for international student assessment (PISA)*. Paris: OECD Publications.

OECD. (2016). *PISA 2015 assessment and analytical framework: Science, reading, mathematic and financial literacy*. Paris: OECD Publishing.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), 943. doi:10.1126/science.aac4716

Ormrod, J. E. (2012). *Human learning* (6. ed.). Upper Saddle River, NJ: Pearson.

Otto, B., & Kistner, S. (2017). Is there a Matthew effect in self-regulated learning and mathematical strategy application? Assessing the effects of a training program with standardized learning diaries. *Learning and Individual Differences, 55*, 75-86. doi:10.1016/j.lindif.2017.03.005

Pai, H.-H., Sears, D. A., & Maeda, Y. (2015). Effects of small-group learning on transfer: A meta-analysis. *Educational Psychology Review, 27*(1), 79-102. doi:10.1007/s10648-014-9260-8

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45*(3), 255-287. doi:10.1037/h0084295

Pan, Y., Liu, H., Lau, P., & Luo, F. (2017). A latent transition analysis of bullying and victimization in Chinese primary school students. *PLoS ONE, 12*(8), e0182802. doi:10.1371/journal.pone.0182802

Park, B., Korbach, A., & Brunken, R. (2015). Do learner characteristics moderate the seductive-details-effect? A cognitive-load-study using eye-tracking. *Journal of Educational Technology & Society, 18*(4), 24-36.

Paulus, W., Nitsche, M. A., & Antal, A. (2016). Application of transcranial electric stimulation (tDCS, tACS, tRNS). *European Psychologist, 21*(1), 4-14. doi:10.1027/1016-9040/a000242

Perkins, D. N., & Salomon, G. (1992). Transfer of learning. In T. Husén & T. N. Postlethwait (Eds.), *International encyclopedia of education* (Vol. 11, pp. 6452-6457). Oxford, England: Pergamon Press.

Petticrew, M., & Roberts, H. (2005). *Systematic reviews in the social sciences: A practical guide*. Oxford: Blackwell.

Pfost, M., Dörfler, T., & Artelt, C. (2012). Reading competence development of poor readers in a German elementary school sample: an empirical examination of the Matthew effect model. *Journal of Research in Reading, 35*(4), 411-426. doi:10.1111/j.1467-9817.2010.01478.x

Piaget, J. (1971). *Biology and knowledge*. Chicago: University of Chicago Press.

Poldrack, R. A., Kittur, A., Kalar, D., Miller, E., Seppa, C., Gil, Y., . . . Bilder, R. M. (2011). The cognitive atlas: Toward a knowledge foundation for cognitive neuroscience. *Frontiers in Neuroinformatics, 5*, 17. doi:10.3389/fninf.2011.00017

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education, 66*(2), 211-227. doi:10.1002/sce.3730660207

Potvin, P., Masson, S., Lafortune, S., & Cyr, G. (2015). Persistence of the intuitive conception that heavier objects sink more: A reaction time study with different levels of interference. *International Journal of Science and Mathematics Education, 13*(1), 21-43. doi:10.1007/s10763-014-9520-6

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research, 42*(1), 185-227. doi:10.1080/00273170701341316

Pressley, M., Wood, E., Woloshyn, V. E., Martin, V., King, A., & Menke, D. (1992). Encouraging mindful use of prior knowledge: Attempting to construct explanatory answers facilitates learning. *Educational Psychologist, 27*(1), 91-109. doi:10.1207/s15326985ep2701_7

Price, A. R., McAdams, H., Grossman, M., & Hamilton, R. H. (2015). A meta-analysis of transcranial direct current stimulation studies examining the reliability of effects on language measures. *Brain Stimulation, 8*(6), 1093-1100. doi:10.1016/j.brs.2015.06.013

Pustejovsky, J. (2017). ClubSandwich: Cluster robust (sandwich) variance estimators with small-sample corrections. R package version 0.2.3.

R Core Team. (2014). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. . Retrieved from http://www.R-project.org/

Raudenbush, S. W. (2009). Analyzing effect sizes: Random-effects models. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (Vol. 2, pp. 295-315). New York: Russell Sage Foundation.

Recht, D. R., & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology, 80*(1), 16. doi:10.1037/0022-0663.80.1.16

Ree, M. J., Carretta, T. R., & Teachout, M. S. (1995). Role of ability and prior knowledge in complex training performance. *Journal of Applied Psychology, 80*(6), 721-730. doi:10.1037/0021-9010.80.6.721

Reeve, C. L., & Hakel, M. D. (2000). Toward an understanding of adult intellectual development: Investigating within-individual convergence of interest and knowledge profiles. *Journal of Applied Psychology, 85*(6), 897-908. doi:10.1037/0021-9010.85.6.897

Renkl, A., Mandl, H., & Gruber, H. (1996). Inert knowledge: Analyses and remedies. *Educational Psychologist, 31,* 115-121. doi:10.1207/s15326985ep3102_3

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

Rittle-Johnson, B., & Schneider, M. (2015). Developing conceptual and procedural knowledge of mathematics. In R. C. Kadosh & A. Dowke (Eds.), *Oxford handbook of numerical cognition* (pp. 1118-1134). Oxford: Oxford University Press.

Rittle-Johnson, B., Schneider, M., & Star, J. (2015). Not a one-way street: Bi-directional relations between procedural and conceptual knowledge of mathematics. *Educational Psychology Review, 27*(4), 587-597. doi:10.1007/s10648-015-9302-x

Rock, D. A. (2007). *A note on gain scores and their interpretation in developmental models designed to measure change in the early school years* (E. T. Service Ed. Vol. 1). Princeton, NJ: Educational Testing Service.

Rosch, E., Mervis, C. B., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*(3), 382-439. doi:10.1016/0010-0285(76)90013-X

Rösler, F., & Heil, M. (2003). The principle of code-specific memory representations. In R. H. Kluwe, G. Lüer, & F. Rösler (Eds.), *Principles of learning and memory* (pp. 71-92). Basel: Birkhäuser.

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005). *Publication bias in meta-analysis: Prevention, assessment and adjustments*. Chichester, England: John Wiley.

Ruiz-Primo, M. A., Briggs, D., Iverson, H., Talbot, R., & Shepard, L. A. (2011). Impact of undergraduate science course innovations on learning. *Science, 331*(6022), 1269-1270. doi:10.1126/science.1198976

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33-58). Hillsdale, NJ: Erlbaum.

Rumelhart, D. E., & Orthony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montaque (Eds.), *Schooling and the acquisition of knowledge* (pp. 99-136). Hillsdale, NJ: Erlbaum.

Russowsky Brunoni, A., & Vanderhasselt, M.-A. (2014). Working memory improvement with non-invasive brain stimulationof the dorsolateral prefrontal cortex: A systematic reviewand meta-analysis. *Brain and Cognition, 86,* 1-9. doi:10.1016/j.bandc.2014.01.008

Sala, G., & Gobet, F. (2017). Does far transfer exist? Negative evidence from chess, music, and working memory training. *Current Directions in Psychological Science, 26*(6), 515-520. doi:10.1177/0963721417712760

Salmeron, L., Kintsch, W., & Canas, J. J. (2006). Reading strategies and prior knowledge in learning from hypertext. *Memory & Cognition, 34*(5), 1157-1171. doi:10.3758/BF03193262

Samson, G. E., Graue, M. E., Weinstein, T., & Wahlberg, H. J. (1984). Academic and occupational performance: A quantitative synthesis. *American Educational Research Journal, 21*(2), 311-321. doi:10.3102/00028312021002311

Sarkar, A., Dowker, A., & Cohen Kadosh, R. (2014). Cognitive enhancement or cognitive cost: Trait-specific outcomes of brain stimulation in the case of mathematics anxiety. *The Journal of Neuroscience, 34*(50), 16605–16610. doi:10.1523/JNEUROSCI.3129-14.2014

Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference Chi-square test statistic. *Psychometrika, 75*(2), 243-248. doi:10.1007/s11336-009-9135-y

Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is it really robust? *Methodology, 6*(4), 147-151. doi:10.1027/1614-2241/a000016

Schmidt, F. L., & Hunter, J. E. (2015). *Methods of meta-analysis: correcting error and bias in research findings* (3rd ed.). Thousand Oaks, CA: Sage.

Schmidt, H. G., Loyens, S. M. M., van Gog, T., & Paas, F. G. W. C. (2007). Problem-based learning is compatible with human cognitive architecture: Commentary on Kirschner, Sweller, and Clark (2006). *Educational Psychologist, 42*(2), 91-97. doi:10.1080/00461520701263350

Schneider, M. (2012). Knowledge integration. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 1684-1686). New York: Springer.

Schneider, M., Beeres, K., Coban, L., Merz, S., Schmidt, S. S., Stricker, J., & De Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science, 20*(3). doi:10.1111/desc.12372

Schneider, M., & Grabner, R. H. (2012). Fact learning. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 1264-1266). New York: Springer.

Schneider, M., & Hardy, I. (2013). Profiles of inconsistent knowledge in children's pathways of conceptual change. *Developmental Psychology, 49*(9), 1639-1649. doi:10.1037/a0030976

Schneider, M., Rittle-Johnson, B., & Star, J. R. (2011). Relations among conceptual knowledge, procedural knowledge, and procedural flexibility in two samples differing in prior knowledge. *Developmental Psychology, 46,* 1525-1538.

Schneider, M., & Stern, E. (2010). The developmental relations between conceptual and procedural knowledge: A multimethod approach. *Developmental Psychology, 46*(1), 178-192. doi:10.1037/a0016701

Schneider, M., Vamvakoussi, X., & Van Dooren, W. (2012). Conceptual change. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning*. New York: Springer.

Schneider, W. (1993). Domain-specific knowledge and memory performance in children. *Educational Psychology Review, 5*, 257-273.

Schneider, W., Körkel, J., & Weinert, F. E. (1989). Domain-specific knowledge and memory performance: A comparison of high- and low-aptitude children. *Journal of Educational Psychology, 81*(3), 306-312. doi:10.1037/0022-0663.81.3.306

Schoenfeld, A. H., & Herrmann, D. J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*(5), 484-494. doi:10.1037/0278-7393.8.5.484

Schrader, F.-W., & Helmke, A. (2015). School achievement: Motivational determinants and processes. In J. Wright (Ed.), *International encyclopedia of the cocial & behavioral sciences* (2nd ed., Vol. 21, pp. 48-54). Oxford: Elsevier.

Schroeders, U., Schipolowski, S., Zettler, I., Golle, J., & Wilhelm, O. (2016). Do the smart get smarter? Development of fluid and crystallized intelligence in 3rd grade. *Intelligence, 59*(November-December 2016), 84-95. doi:10.1016/j.intell.2016.08.003

Schuler, H., Funke, U., & Baron-Boldt, J. (1990). Predictive validity of school grades - a meta-analysis. *Applied Psychology: An International Review, 39*(1), 89-103.

Schwartz, D. L., Varma, S., & Martin, L. (2008). Dynamic transfer and innovation. In S. Vosniadou (Ed.), *Handbook of conceptual change* (pp. 479-506). Mahwah: Erlbaum.

Schwartz, M., & Fischer, K. W. (2004). Building general knowledge and skill: Cognition and microdevelopment in science learning. In A. Demetriou & A. Raftopoulos (Eds.), *Cognitive developmental change: Theories, models, and measurement* (pp. 157-185). Cambridge, UK: Cambridge University Press.

Science Europe. (2012). *Horizon 2020: Excellence Counts*. Retrieved from http://www.scienceeurope.org/wp-content/uploads/2014/05/SE_H2020_Excellence_Counts_FIN.pdf

Sharp, A. C., Sinatra, G. M., & Reynolds, R. E. (2008). The development of children's orthographic knowledge: A microgenetic perspective. *Reading Research Quarterly, 43*(3), 206-226. doi:10.1598/RRQ.43.3.1

Shea, B. J., Hamel, C., Wells, G. A., Bouter, L. M., Kristjansson, E., Grimshaw, J., . . . M., B. (2009). AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *Journal of Clinical Epidemiology, 62*(10), 1013-1020. doi:10.1016/j.jclinepi.2008.10.009

Shing, Y. L., & Brod, G. (2016). Effects of prior knowledge on memory: Implications for education. *Mind, Brain, and Education, 10*(3), 153-161. doi:10.1111/mbe.12110

Shtulman, A., & Harrington, K. (2016). Tensions between science and intuition across the lifespan. *Topics in Cognitive Science, 8*, 118-137. doi:10.1111/tops.12174

Shtulman, A., & Valcarcel, J. (2012). Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition, 124*(2), 209-215. doi:10.1016/j.cognition.2012.04.005

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.

Siegler, R. S., Fazio, L. K., Bailey, D. H., & Zhou, X. (2013). Fractions: The new frontier for theories of numerical development. *Trends in Cognitive Sciences, 17*(1), 13-19. doi:10.1016/j.tics.2012.11.004

Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard University Press.

Smith III., J. P., diSessa, A. A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences, 3*(2), 115-163. doi:10.1207/s15327809jls0302_1

Snow, R. E. (1989). Aptitude, instruction, and individual development. *International Journal of Educational Research, 13*(8), 869-881. doi:10.1016/0883-0355(89)90070-0

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science, 10*(1), 89-96. doi:10.1111/j.1467-7687.2007.00569.x

Stagg, C. J., Best, J. G., Stephenson, M. C., O'Shea, J., Wylezinska, M., Kincses, Z. T., . . . Johansen-Berg, H. (2009). Polarity-sensitive modulation of cortical neurotransmitters by transcranial stimulation. *Journal of Neuroscience, 29*(16), 5202-5206. doi:10.1523/JNEUROSCI.4432-08.2009

Stagg, C. J., & Nitsche, M. A. (2011). Physiological basis of transcranial direct current stimulation. *The Neuroscientist, 17*(1), 37-53. doi:10.1177/1073858410386614

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy *Reading Research Quarterly, 21*(4), 360-407. doi:10.1177/0022057409189001-204

Stein, M. K., & Smith, M. S. (1998). Mathematical tasks as a framework for reflection: From research to practice. *Mathematics Teaching in the Middle School, 3*, 268-275.

Steinberg, L. (2014). *Adolescence* (10th Ed.). New York: McGraw-Hill.

Steinmayr, R., Meißner, A., Weidinger, A. F., & Wirthwein, L. (2014). Academic achievement. *Oxford Bibliographies.* Retrieved from http://www.oxfordbibliographies.com/view/document/obo-9780199756810/obo-9780199756810-0108.xml

Stern, E. (2001). Intelligence, prior knowledge, and learning. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (Vol. 11, pp. 7670-7674). Oxford: Elsevier Science.

Stern, E. (2017). Individual differences in the learning potential of human beings. *Science of Learning, 2*(2), 1-7. doi:10.1038/s41539-016-0003-0

Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251-296. doi:10.1023/a:1022193728205

Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics* (6th ed.). Essex: Pearson Education Limited.

Tanaka, D., Kiyokawa, S., Yamada, A., Dienes, Z., & Shigemasu, K. (2008). Role of selective attention in artificial grammar learning. *Psychonomic Bulletin & Review, 15*(6), 1154-1159. doi:10.3758/PBR.15.6.1154

Tandon, P. N., & Singh, N. C. (2016). Educational neuroscience: Challenges and opportunities *Annals of Neuroscience, 23*, 63-65. doi:10.1159/000443560

Tanner-Smith, E. E., & Tipton, E. (2014). Robust variance estimation with dependent effect sizes: Practical considerations and a software tutorial in Stata and SPSS. *Research Synthesis methods, 5*(1), 13-30. doi:10.1002/jrsm.1091

Tanner-Smith, E. E., Tipton, E., & Polanin, J. R. (2016). Handling complex meta-analytic data structures using robust variance estimates: A tutorial in R. *Journal of Developmental and Life-Course Criminology, 2*(1), 85-112. doi:10.1007/s40865-016-0026-5

Tanner-Smith, E. E., & Wilson, S. J. (2013). A meta-analysis of the effects of dropout prevention programs on school absenteeism. *Prevention Science, 14*(5), 468-478. doi:10.1007/s11121-012-0330-1

Thompson, R. A., & Zamboanga, B. L. (2003). Prior knowledge and its relevance to student achievement in introduction to psychology. *Teaching of Psychology, 30*(2), 96-101. doi:10.1207/S15328023TOP3002_02

Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions II: The estimation of magnitudes. *Psychological Review, 8*(4), 384-395. doi:10.1037/h0071280

Trapmann, S., Hell, B., Weigand, S., & Schuler, H. (2007). Die Validität von Schulnoten zur Vorhersage des Studienerfolgs – eine Metaanalyse. *Zeitschrift für pädagogische Psychologie, 21*(1), 11-27. doi:10.1024/1010-0652.21.1.11

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

Tynjälä, P. (1999). Towards expert knowledge? A comparison between a constructivist and a traditional learning environment in the university. *International Journal of Educational Research, 31*(5), 357-442. doi:10.1016/S0883-0355(99)00012-9

Ullén, F., Hambrick, D. Z., & Mosing, M. A. (2016). Rethinking expertise: A multifactorial gene-environment interaction model of expert performance. *Psychological Bulletin, 142*(4), 427-446. doi:10.1037/bul0000033

van Elk, M., Matzke, D., Gronau, Q. F., Guan, M., Vandekerckhove, J., & Wagenmakers, E.-J. (2015). Meta-analyses are no substitute for registered replications: A skeptical perspective on religious priming. *Frontiers in Psychology, 15*(6), 1365. doi:10.3389/fpsyg.2015.01365

van Kesteren, M. T. R., Rijpkema, M., Ruiter, D. J., & Fernández, G. (2010). Retrieval of associative information congruent with prior knowledge is related to increased medial prefrontal activity and connectivity. *The Journal of Neuroscience, 30*(47), 15888-15894. doi:10.1523/JNEUROSCI.2674-10.2010

van Kesteren, M. T. R., Rijpkema, M., Ruiter, D. J., Morris, R. G. M., & Fernández, G. (2014). Building on prior knowledge: Schema-dependent encoding processes relate to academic performance. *Journal of Cognitive Neuroscience, 26*(10), 2250-2261. doi:10.1162/jocn_a_00630

van Loon, M. H., de Bruin, A. B. H., van Gog, T., & van Merriënboer, J. J. G. (2013). Activation of inaccurate prior knowledge affects primary-school students'

metacognitive judgments and calibration. *Learning and Instruction, 24*(April 2013), 15-25. doi:10.1016/j.learninstruc.2012.08.005

Vandenbrouke, L., Verschueren, K., & Baeyens, D. (2017). The development of executive functioning across the transition to first grade and its predictive value for academic achievement. *Learning and Instruction, 49*, 103-112. doi:10.1016/j.learninstruc.2016.12.008

Vickers, A. J. (2001). The use of percentage change from baseline as an outcome in a controlled trial is statistically inefficient: A simulation study. *BMC Medical Research Methodology, 1*(1), 6. doi:10.1186/1471-2288-1-6

Viechtbauer, W. (2010). Conducting meta-analysis in R with the metafor package. *Journal of Statistical Software, 36*(3), 1-48.

Viechtbauer, W., & Cheung, M. W.-L. (2010). Outlier and influence diagnostics for meta-analysis. *Reserach Synthesis methods, 1*(2), 112-125. doi:10.1002/jrsm.11

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4*(1), 45-69. doi:10.1016/0959-4752(94)90018-3

Vosniadou, S. (2008). *International handbook for research on conceptual change*. New York: Routledge.

Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology, 24*(4), 535-585.

Vosniadou, S., Ioannides, C., Dimitrakopoulou, A., & Papademetriou, E. (2001). Designing learning environments to promote conceptual change in science. *Learning and Instruction, 11*(4-5), 381-419.

Vosniadou, S., & Skopeliti, I. (2017). Is it the earth that turns or the sun that goes behind the mountains? Students' misconceptions about the day/night cycle after reading a science text. *International Journal of Science Education, 39*(15), 2027-20521. doi:10.1080/09500693.2017.1361557

Vosniadou, S., Vamvakoussi, X., & Skopeliti, I. (2008). The framework theory approach to the problem of conceptual change. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 3-34). New York: Routledge.

Vygotsky, L. S. (1978). *Mind and society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.

Walberg, H. J., & Tsai, S.-L. (1983). Matthew effects in education. *American Educational Research Journal, 20*(3), 359-373. doi:10.3102/00028312020003359

Waltz, J. A., Lau, A., Grewal, S. K., & Holyoak, K. J. (2000). The role of working memory in analogical mapping. *Memory & Cognition, 28*(7), 1205-1212. doi:10.3758/BF03211821

Wang, S.-H., & Morris, R. G. M. (2010). Hippocampal-neocortical interactions in memory formation, consolidation, and reconsolidation. *Annual Review of Psychology, 61*, 49–79. doi:10.1146/annurev.psych.093008.100523

Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What's past is prologue: Relations between early mathematics knowledge and high school achievement. *Educational Researcher, 43*(7), 352-350. doi:10.3102/0013189X14553660

Weil, L. G., Fleming, S. M., Dumontheil, I., Kilford, E. J., Weil, R. S., Rees, G., . . . Blakemore, S.-J. (2013). The development of metacognitive ability in adolescence. *Consciousness and Cognition, 22*, 264-271. doi:10.1016/j.concog.2013.01.004

Weinert, F. E. (1999). Definition and selection of competencies. Concepts of competence. Retrieved from https://pdfs.semanticscholar.org/8b88/efa9dd5e0a4b605aea6e5e3b9ec640beb089.pdf

Weiser, M., & Shertz, J. (1983). Programming problem representation in novice and expert programmers. *International Journal of Man-Machine Studies, 19*(4), 391-398. doi:10.1016/S0020-7373(83)80061-3

Wellman, H. M., & Gelman, S. A. (1992). Cognitive development: Foundational theories of core domains. *Annual Review of Psychology, 43*(1), 337-375. doi:10.1146/annurev.ps.43.020192.002005

Williams, R. H., & Zimmerman, D. W. (1996). Are simple gain scores obsolete? *Applied Psychological Measurement, 20*(1), 59-69. doi:10.1177/014662169602000106

Wolpert, R. S. (1990). Recognition of melody, harmonic accompaniment, and instrumentation: Musicians vs. nonmusicians. *Music Perception: An Interdisciplinary Journal, 8*(1), 95-105. doi:10.2307/40285487

Woltz, D. J., Gardner, M. K., & Bell, B. G. (2000). Negative transfer errors in sequential cognitive skills: Strong-but-wrong sequence application. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*(3), 601-625. doi:10.1037/0278-7393.26.3.601

Wu, R., & Zhao, J. (2017). Prior knowledge of object associations shapes attentional templates and information acquisition. *Frontiers in Psychology, 8,* 843. doi:10.3389/fpsyg.2017.00843

Wurzman, R., Hamilton, R. H., Pascual-Leone, A., & Fox, M. D. (2016). An open letter concerning do-it-yourself users of transcranial direct current stimulation. *Annals of Neurology, 80,* 1-4. doi:10.1002/ana.24689

Yates, G. C. R., & Chandler, M. (1991). The cognitive psychology of knowledge: Basic research findings and educational implications. *Australian Journal of Education, 35*(2), 131-153. doi:10.1177/000494419103500203

Yu, C., Zhong, Y., & Fricker, D. (2012). Selective attention in cross-situational statistical learning: Evidence from eyetracking. *Frontiers in Psychology, 3,* Article 148. doi:10.3389/fpsyg.2012.00148

Zacharia, Z. C., Loizou, E., & Papaevripidou, M. (2012). Is physicality an important aspect of learning through science experimentation among kindergarten students? *Early Childhood Research Quartely, 27*(3), 447-457. doi:10. 1O16/j.ccrcsq.20 12.02.004

Zimmermann, D. W., & Williams, R. H. (1998). Reliability of gain scores under realistic assumptions about properties of pre-test and post-test scores. *Mathematical and Statistical Psychology, 51*, 343-351. doi:10.1111/j.2044-8317.1998.tb00685.x

Authorship and Publication Status

Chapters 5 to 7 present research articles that are or will soon be submitted for publication. Hereinafter, the authors and publications statuses of the papers are enlisted.

1. Study 1 (Chapter 5): **Simonsmeier, B. A.**, Grabner, R. H., Hein, J., Krenz, U., Schneider, M. (2018). Electrical brain stimulation (tES) improves learning more than performance: A meta-analysis. *Neuroscience & Biobehavioral Reviews, 84*, 171-181. doi: 10.1016/j.neubiorev.2017.11.001

2. Study 2 (Chapter 6): **Simonsmeier, B. A.**, Flaig, M., Deiglmayr, A., Schalk, L., & Schneider, M. (manuscript in preparation). Domain-specific prior knowledge and learning: A meta-analysis.

3. Study 3 (Chapter 7): Flaig, M., **Simonsmeier, B. A.**, Mayer, A.-K., Rosman, T., Gorges, J., & Schneider, M. (in press). Conceptual change and knowledge integration as learning processes in higher education: A latent transition analysis. *Learning and Individual Differences*.

Declaration of Authorship

I hereby declare that the thesis submitted is my own unaided work. All direct or indirect sources used are acknowledged as references. This dissertation was not previously presented to another examination board in order to obtain an academic title.

Eidesstattliche Erklärung

Hiermit erkläre ich, dass die vorliegende Dissertationsschrift von mir selbständig angefertigt wurde und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet wurden. Zudem wurde die Arbeit an keiner anderen Universität zur Erlangung eines akademischen Grades eingereicht.

Trier, _____ _____

Date/Datum Bianca A. Simonsmeier (MSc)