Trier University

Faculty I - Department of Psychology

Knowledge Acquisition and Transfer – Three Meta-Analytic Investigations

Dissertation submitted in fulfillment of the requirements for the degree of

Doctor rerum naturalium (Dr. rer. nat.) in Psychology at Trier University

by

Aline Chantal Dlugosch

Supervised by

Prof. Dr. Michael Schneider

Prof. Dr. Nicola Baumann

Submitted on July 14th 2023

Place of Dissertation: Trier, Germany

Place of Dissertation / Ort der Dissertation: Trier, Germany

Acknowledgments

In the following, I would like to thank certain people that accompanied me through the last months of work on this dissertation. First, I would like to mention those who worked together with me in the Department of Educational Psychology at Trier University. I am grateful for the continuous scientific support and guidance of my advisor Prof. Dr. Michael Schneider. I learned a lot in the last three years, both scientifically and humanly. I am also thankful that Prof. Dr. Nicola Baumann agreed to supervise my work in this dissertation as the second supervisor. Thanks to my colleagues who always supported me and my work scientifically as well as emotionally, Claire Forsmann, Thomas Simacek, Bianca Simonsmeier-Martin, and Uschi Weiland. Similarly, the research assistants Jonathan, Marlene, Moritz, Romina, and Yannick contributed to this work by providing their assistance throughout the last months.

I want to express deep gratitude to my family for their everlasting support. First and foremost, my parents Gabi and Winfried, and their partners, Ricarda and Dieter, have spared no time and effort to support me and my work. They each encouraged me in their own way: From horizon-expanding content talks to relaxing walks with the dogs, they provided me with everything I needed to achieve my goals.

I want to thank my partner Marcel, who was my safe haven through the journey of the last months and my calming influence in stressful times. He always shows me what I can do, even when I might doubt myself.

I cannot put into words how grateful I am.

Abstract

Knowledge acquisition comprises various processes. Each of those has its dedicated research domain. Two examples are the relations between knowledge types and the influences of person-related variables. Furthermore, the transfer of knowledge is another crucial domain in educational research. I investigated these three processes through secondary analyses in this dissertation. Secondary analyses comply with the broadness of each field and yield the possibility of more general interpretations. The dissertation includes three meta-analyses: The first meta-analysis reports findings on the predictive relations between conceptual and procedural knowledge in mathematics in a cross-lagged panel model. The second meta-analysis focuses on the mediating effects of motivational constructs on the relationship between prior knowledge and knowledge after learning. The third meta-analysis deals with the effect of instructional methods in transfer interventions on knowledge transfer in school students. These three studies provide insights into the determinants and processes of knowledge acquisition and transfer. Knowledge types are interrelated; motivation mediates the relation between prior and later knowledge, and interventions influence knowledge transfer. The results are discussed by examining six key insights that build upon the three studies. Additionally, practical implications, as well as methodological and content-related ideas for further research, are provided.

Table of Contents

List of Figures

List of Tables

1 Introduction

Knowledge is a widely researched construct but lacks a clear definition (e.g., Pritchard, 2013). However, researchers and practitioners across many domains agree on its importance for success in education and beyond (e.g., knowledge-is-power hypothesis; Greve et al., 2019; Möhring et al., 2018). Different processes belonging to the construct of knowledge are inspected in several research areas. Some examples are the relations between different knowledge types (e.g., Rittle-Johnson & Alibali, 1999), motivational factors influencing knowledge acquisition (e.g., Vu et al., 2022), and interventions to foster knowledge transfer to other contexts or tasks (e.g., Kaminske et al., 2020). A broad range of past and current research exists for these exemplary fields.

Questions about knowledge have concerned researchers since the beginning of philosophy. Starting with epistemology as the study of knowledge and knowing, issues about the nature and form of human knowledge have been proposed (Buehl & Alexander, 2001). Additionally, Plato tried to explore these elements through dialogues (e.g., Gulley, 2013), and Kant (1958) studied additional parameters by differentiating rational and empirical reasoning. The investigations took a psychological turn in the twentieth century regarding the relationship between knowledge and education (e.g., Dewey, 1916; Whitehead, 1967). Up to this point, the general importance of knowledge remains clear. However, none of the proposed definitions are widely acknowledged in all research domains.

Not only researchers are interested in the topic of knowledge. Also, the public demands insight into the concept of knowledge, mainly in the setting of education. The PISA studies are one famous example in which the school students' knowledge is assessed and compared across countries in different domains. The results are essential not only for researchers but likewise for practitioners, parents, and institutions. In PISA 2015, Germany scored 509 points in the STEM domains, which is a result significantly above the average

(Schiepe-Tiska, Rönnebeck, et al., 2015). The PISA studies did not only capture aspects of knowledge: Motivation and attitudes were assessed in addition to conceptual, procedural, and epistemic knowledge (Schiepe-Tiska, Simm, et al., 2015). This underlines the further importance of other personal factors in combination with success in knowledge acquisition.

The current dissertation gives an overview of three facets of processes associated with knowledge in the form of three meta-analyses. These are the longitudinal relations between knowledge types, the mediation of the relationship between prior knowledge and knowledge after learning, as well as instructional influences on knowledge transfer. The dissertation is divided into three main parts: It starts with the theoretical background of knowledge acquisition and transfer (Chapters 2 to 4), then presents three studies as the core section (Chapters 5 to 7). The three studies offer insights into the predictive relations between conceptual and procedural knowledge in mathematics (Study 1 in Chapter 5), the relevance of motivational mediators on the relationship between prior knowledge and knowledge after learning (Study 2 in Chapter 6), and the impact of instructional methods in interventions on knowledge transfer in school students (Study 3 in Chapter 7). Finally, the dissertation closes with a general discussion and conclusion (Chapter 8).

2 Knowledge

Knowledge plays a central role in the research on learning and education (de Jong & Ferguson-Hessler, 1996). Despite its essential part in past and current investigations, there is no clear definition of the construct *knowledge*. An issue of defining knowledge is the question of where to start: One could begin by looking at individual cases and their similarities, but if one does not already know what knowledge is or how it is defined, how can it be correctly identified through single cases (Pritchard, 2013)? Early traditional philosophers like Plato and Aristotle developed epistemology as a theory of knowledge by trying to define it (Buehl & Alexander, 2001). Today, however, their responses are not generally accepted (Neta & Pritchard, 2009; Russell, 1972).

Definitions of knowledge differ in their level of analysis (Phye, 1997). From a reductionist view, knowledge can be investigated at the level of the central nervous system using neurological and biopsychological explanations (e.g., Brown & Smith, 2003; Sayer, 2010). The other extreme includes explaining knowledge on the dimension of metaphysics through philosophical thoughts about reality (e.g., Hossack, 2007; Phye, 1997; Simon, 2020).

Another possible distinction is the differentiation between rationalism and empiricism (Bolisani & Bratianu, 2018; Longworth, 2009). In rationalism, knowledge results from a reasoning process called rational reasoning (e.g., Descartes, 1997; Gorham, 2002). Empiricism is the opposite perspective, including the connection of ideas and sensory information. According to this, knowledge is created through sensory interfaces with the world (Bolisani & Bratianu, 2018; Keil et al., 1998).

Knowledge is seen as a "justified true belief" (Nonaka & Takeuchi, 1995, p. 87) in the field of epistemology. It incorporates three basic conditions: The truth, the belief, and the justification condition. The truth condition states that the proposition a person knows needs to be true, which differentiates an opinion from knowledge. Second, the person must believe in the proposition, called the belief condition. Last, the justification condition proposes that knowledge needs a practical way of justifying the proposition's truth (Ayer, 2009; Bolisani & Bratianu, 2018; Gettier, 2009; Neta & Pritchard, 2009).

Spelke and Kinzler (2007) present a more biological view of knowledge: They believe all humans have several separable core knowledge systems. Studies on knowledge's ontogenetic and phylogenetic origins concentrated on four of these core systems (Spelke, 2004): objects, actions, numbers, and places. The most extensively studied of these is the system of object representation (Spelke & Kinzler, 2007). Knowledge in this system enables human infants to perceive object boundaries and movements (e.g., Valenza et al., 2006). New systems, consisting of new skills or beliefs, are built on these foundations.

These philosophical and biological dimensions of knowledge above are helpful in theoretical discourses but do not provide information for instructors or teachers when collaborating with their students. To reach this goal, a definition of knowledge should imply some correspondence with the learner and their perspective. This fits the cognitive point of view: According to cognitive researchers, knowledge includes "an individual's personal stock of information, skills, experiences, beliefs, and memories" (Alexander et al., 1991, p. 317). Thus, knowledge represents a person's learning history and must not be verified as valid.

There is no general definition of the broad concept of knowledge. In the research area of educational psychology, the cognitive definition yields the most advantages. It incorporates the student's learning history and possible correspondences between instructors and learners. Hence, this definition will be used as a basis for this dissertation.

Moreover, knowledge can be distinguished based on several aspects. First, knowledge usually is divided into different knowledge domains. A knowledge domain includes a subset of knowledge in a specific content area (Alexander et al., 1991). This knowledge is restricted to a specific field or thought (e.g., Voss et al., 1986). The classification of knowledge in knowledge domains allows more thorough investigations in that research area. Furthermore, it is possible to differentiate between knowledge types (see Chapter 2.1).

2.1 Knowledge Types

Knowledge can be categorized into several knowledge types and qualities (de Jong & Ferguson-Hessler, 1996). Types of knowledge focus on specific characteristics in which knowledge differs. The quality of knowledge encompasses the associated properties knowledge can have. Considering these two concepts, it is feasible to create a matrix that structures general topics, for example, learning goals (de Jong & Ferguson-Hessler, 1996). This chapter focuses on three types of knowledge proposed by several theoretical perspectives (e.g., de Jong & Ferguson-Hessler, 1996; Rittle-Johnson et al., 2001): declarative knowledge, conceptual knowledge, and procedural knowledge.

2.1.1 Declarative Knowledge

Declarative knowledge, or propositional knowledge (Pritchard, 2013), contains information on a person’s environment, for example, what learners can see or hear, and lets them recall past events (Ten Berge & Hezwijk, 1999). Declarative knowledge also includes concepts, ideas, principles, and theories (Ohlsson, 1994, 1996). One possible distinction within declarative knowledge is the difference between semantic and episodic memory. This distinguishes the memory of general knowledge of the world independent of life experiences from such life experiences (Bauer, 2006; Broadbent, 1989; Tulving, 1993). Using words like *remembering*, which points to episodic events, and *knowing* about semantic facts, illustrates this distinction. To sum up, declarative knowledge entails information about events and facts (e.g., Kump et al., 2015; Ten Berge & Van Hezewijk, 1999) that can be verbalized instantly (Neely, 1989).

The structure of declarative knowledge can be represented as a semantic network (Chi & Ohlsson, 2005). This network consists of nodes and links that depict concepts and their relations (Chi & Ohlsson, 2005; Hartley & Barnden, 1997). These connections are grouped by domain so that similar pieces of knowledge cluster together (Sowa, 1984, 2014). It is possible to explore such networks through computer simulations (Abelson, 1973; Quillian, 1968; Rips & Medin, 2005).

2.1.2 Conceptual Knowledge

Knowledge about concepts is called *conceptual knowledge*. There are different views on whether this knowledge type is a part of declarative knowledge (e.g., Schneider & Stern, 2010) or an individual knowledge type (e.g., Hiebert & Lefevre, 1986). Conceptual knowledge includes static knowledge about facts, concepts, and principles that apply within a knowledge domain (de Jong & Ferguson-Hessler, 1996). Several definitions of conceptual knowledge exist, depending on the knowledge domain. In the following, some of them are reviewed for the knowledge domain of mathematics (Crooks & Alibali, 2014). In this domain, conceptual knowledge is described as knowledge of connections (e.g., Hiebert & Lefevre, 1986), knowledge of general principles (e.g., de Jong & Ferguson-Hessler, 1996), knowledge of underlying principles (e.g., Baroody et al., 2007), knowledge of organizing categories (e.g., Byrnes, 1992), knowledge of symbol meanings (e.g., Ploger & Hecht, 2009), or knowledge of domain structures (e.g., Robinson & Dube, 2009).

Conceptual knowledge is also assumed to be stored mentally in relational representations (Schneider & Stern, 2010). Punctate and holographic models are examples of more precise models of this representation in the human brain and the neural connections between concepts (e.g., Feldman, 1986). Punctate models purport that each concept is represented by one neuron (Barlow, 1972) and that a category-based memory network depicts their interrelations (Shastri, 1985). On the other hand, holographic models represent concepts and interrelations in the form of a linear threshold matrix (Feldman, 1986).

2.1.3 Procedural Knowledge

Procedural knowledge is defined as the knowledge of procedures or simply as *knowing how* (Byrnes & Wasik, 1991; Rittle-Johnson et al., 2001). A procedure includes actions that are carried out to accomplish a goal. Subsequently, procedural knowledge differs from other forms of knowledge because of procedures' innate sequential and goal-directed nature (Hiebert & Lefevre, 1986). Procedural knowledge can furthermore be described as the knowledge of operators and the conditions under which these operators are applied to reach an intention (Anderson, 1993; Baroody, 2003; Rittle-Johnson et al., 2001; Schneider & Stern, 2010). It involves a combination of physical and mental activities, ranging from simple to complex actions (Ormrod, 2012). Such complex procedures are acquired slowly over some time (e.g., Ericsson, 2003).

Based on prominent definitions (e.g., Hiebert & Lefevre, 1986), procedural knowledge is characterized as superficial and poor in connections (Star, 2005). These definitions miss the variety of procedures and their different qualities of connections (Anderson, 1982). Therefore, Star (2005) reconceptualized procedural knowledge into superficial and deep procedural knowledge. Superficial procedural knowledge corresponds to the common usage of procedural knowledge, whereas deep procedural knowledge is associated with comprehension, flexibility, and critical judgment (Star, 2005). This deep procedural knowledge can be measured by letting learners explain the steps in a procedure to achieve the goal (Star, 2007).

2.1.4 Interrelations Between Knowledge Types

The knowledge types are difficult to measure separately because of their interrelations (Baroody et al., 2007; Star & Stylianidis, 2013). Additionally, there are no established standards for measuring single knowledge types. Recent evidence showed that conceptual and procedural fraction knowledge are empirically separable when using appropriate measures (Bempeni et al., 2018; Lenz et al., 2020). However, most studies do not use such exclusively developed instruments. This makes it difficult to measure knowledge types validly and independently from each other (Schneider & Stern, 2010).

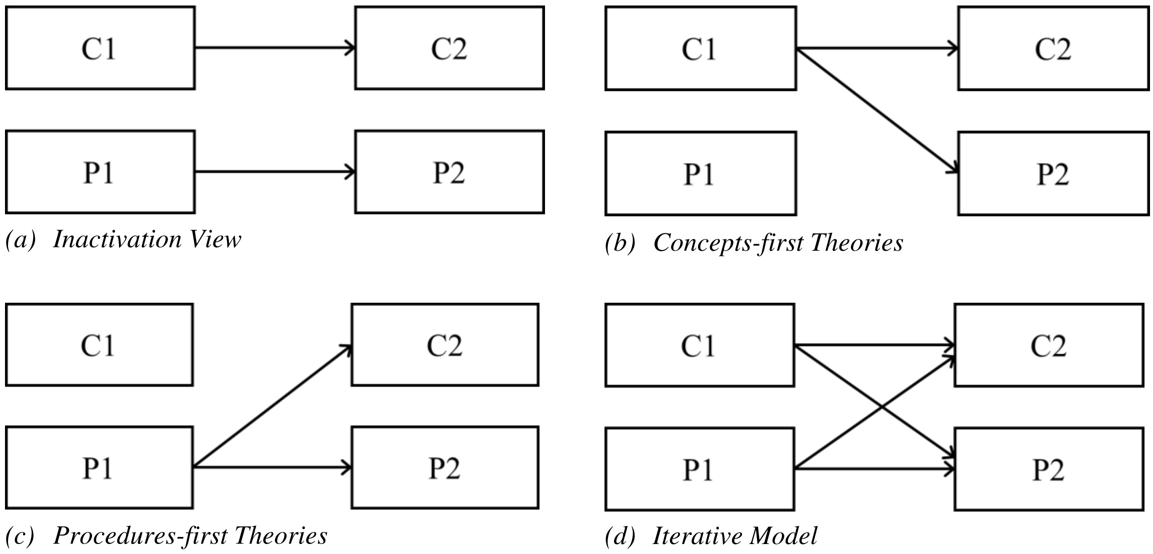
There are several theoretical views on the relations between conceptual and procedural knowledge. Four theoretical viewpoints on these longitudinal relations are traditionally mentioned (see Figure 1): The *Inactivation View* (Haapasalo & Kadjievich, 2000) states that the two knowledge types develop independently from each other (e.g., Nesher, 1986; Resnick & Omansun, 1987). *Concept-first Theories* predicate that learners develop conceptual knowledge first. Subsequently, procedural knowledge derives from it through unidirectional construction and evaluation of problem-solving processes (e.g., Donlan et al., 2007; Geary, 1994; Gelman, 1993; Gelman & Williams, 1998). However, *Procedures-first Theories* (Weaver et al., 2018) state that learners develop procedural knowledge first and then build conceptual knowledge on this basis through unidirectional abstraction processes (e.g., Baroody & Gannon, 1984; Baroody & Ginsburg, 1986; Briars & Siegler, 1984; Kerslake, 1986; Siegler & Stern, 1998). The fourth theoretical viewpoint is called the *Iterative Model* (Rittle-Johnson et al., 2001). According to this viewpoint, the causal relations between conceptual and procedural knowledge are bidirectional. Therefore, the increase in one of them leads to a subsequent increase in the other (e.g., Canobi, 2009; Hecht & Vagi, 2012; Rittle-Johnson & Koedinger, 2009; Schneider et al., 2011).

2.2 Determinants of Knowledge Acquisition

Knowledge acquisition, often synonymously called *learning*, is one of the most crucial goals in educational settings. There are several theories of knowledge acquisition (Renkl, 2009). To study the aspects and characteristics of knowledge, it is essential to look at how learners acquire knowledge. For example, the learner’s prior knowledge determines later knowledge acquisition (Chapter 2.2.1), and motivational factors can further affect it (Chapter 2.2.2).

Figure 1

Four Theoretical Viewpoints on Longitudinal Relations Between Conceptual and Procedural Knowledge



Note. The four models a, b, c, and d show the proposed influences of conceptual knowledge (C1) and procedural knowledge (P1) at measurement point one on both knowledge types at a later measurement point (C2, P2).

2.2.1 Prior Knowledge

Learners can only connect new information if they already have gained knowledge about the learning contents (Ormrod, 2012). This *prior knowledge* influences a learner's ability to encode all kinds of information. Prior knowledge is defined as the sum of knowledge an individual already possesses (Alexander et al., 1991). Learners always build on prior knowledge to learn new content (Schwartz et al., 2007). Hence, learners with a large body of information in long-term memory can generate more ideas relating to their new experiences (Ormrod, 2012). Learners with insufficient relevant knowledge use inefficient rote-learning strategies instead. Several overlapping theories propose processes in which prior knowledge affects learning (Dochy, 1994). Example key processes are the elaboration of redundant retrieval paths or the greater availability of information during the learning process (Dochy, 1994). Another two common approaches include the role of prior knowledge with a conceptual versus a procedural point of view (Schwartz et al., 2007): From the conceptual point of view, the impact of prior knowledge builds upon students' intuitions and experiences. Focusing on procedural prior knowledge, the effect builds upon the mastery of component skills and their combinations (Schwartz et al., 2007).

Numerous studies have illustrated the influence of prior knowledge (e.g., Anderson, 1981; Hall, 1989; Ormrod et al., 1988; Schneider, 1993). It is the most important factor for long-term memory storage (Haskell, 2001). Prior declarative and predominantly procedural knowledge from previous courses influenced pharmacy students' later achievement (Hailikari et al., 2008). Pazzani (1991) found effects of prior knowledge on concept acquisition. Additionally, a recent meta-analysis inspected the relationship between prior knowledge and knowledge acquisition based on 8776 effect sizes (Simonsmeier et al., 2022): The meta-analysis showed that relative differences in knowledge between individuals were highly stable from before to after learning. This makes prior knowledge a suitable predictor for later knowledge. However, the included primary studies differed strongly in their effect sizes, resulting in a broad prediction interval around zero. Based on this result, the discussion about the impact of prior knowledge on knowledge acquisition must resume.

Recently, in the domain of text comprehension, prior knowledge was described as having four dimensions: amount, accuracy, specificity, and coherence (McCarthy & McNamara, 2021). This implies that the influence of prior knowledge depends on more than just the amount a learner possesses (Brod, 2021). One example is the activation of relevant prior knowledge congruent with new information (e.g., Bein et al., 2015; Bransford & Johnson, 1972; Brod & Shing, 2019; McWeeny et al., 1987). On the other hand, the activation of irrelevant and incongruent prior knowledge hinders further learning (Castel et al., 2007; Chi, 2008).

Besides correct information, prior knowledge can include errors or misconceptions as well. According to diSessa (2006), misconceptions are students' persistent false beliefs despite being contradicted by established evidence (Bensley & Lilienfeld, 2017). These misconceptions often arise from everyday sources. Several studies have proven misconceptions problematic (e.g., Eaton et al., 1984; Kendeou & van den Broek, 2005). If learners use inaccurate knowledge to encode new information that is inconsistent with their existing knowledge, they might either ignore the new information or distort it to make it consistent (Ormrod, 2012). Therefore, misconceptions might be more detrimental than knowing nothing about a topic. Prior knowledge is mostly stable in terms of misconceptions (e.g., Carey, 1986). However, it is not entirely resistant to change. It is possible to alter misconceptions through explicit activation and confronting with contrary evidence (Bensley & Lilienfeld, 2017; Kowalski & Taylor, 2009). On the other hand, new misinformation can also negatively influence correct prior knowledge (Fazio et al., 2013).

Prior knowledge is among the most influential predictors of knowledge acquisition. However, there is no clear evidence of how strong prior knowledge affects learning and whether there are other determinants. For example, motivational factors in a learner might additionally play a role in the relationship between prior knowledge and learning outcomes.

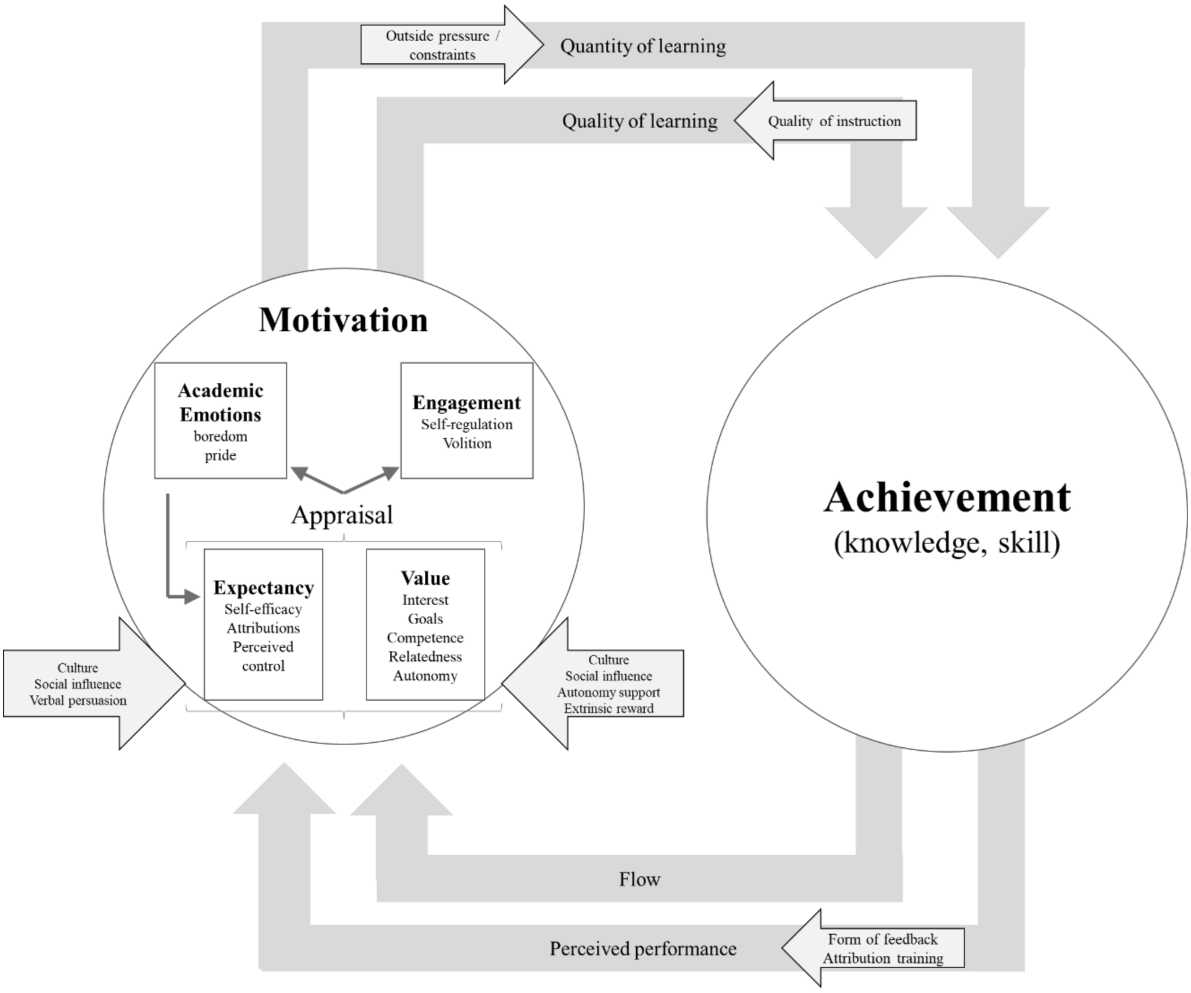
2.2.2 Motivation

Motivation is often researched in educational psychology (Koenka, 2020; Urhahne & Wijnia, 2023), focusing on its influence on behavior, learning, and performance (Schunk et al., 2014). A general model of the relationship between motivation and achievement is presented by Vu and colleagues (2022), namely the *Motivation-Achievement Cycle* (see Figure 2). In this model, motivation and achievement affect each other through behavioral intermediaries and achievement-related constructs with the inclusion of outside influences. Achievement is operationalized through the acquisition of knowledge or skill. Motivation is mainly used as an umbrella term for several unique constructs, each with its own theoretical and empirical basis.

Urhahne and Wijnia (2023) integrated the essential theories into a basic model of motivation in education, including the process from the interaction between the situation and the self to the consequences of their actions. To illustrate the variety, the constructs of interest, self-concept, self-efficacy, and intrinsic as well as extrinsic motivation are described in the following.

Figure 2

Motivation-Achievement Cycle (Vu et al., 2022)



Note. The *Motivation-Achievement Cycle* (Vu et al., 2022) includes motivation-achievement interactions. In the circle on the left are motivation constructs. The right circle and the arrows from right to left are achievement constructs. The arrows from left to right include behavioral intermediaries. Light gray arrows denote outside influences (own representation, based on Vu et al., 2022, p. 41).

2.2.2.1 Interest. *Interest* is a content-specific motivational construct that includes intrinsic feelings- and value-related valences (Schiefele, 1991). It contains affective and cognitive components (Hidi & Renninger, 2006). Additionally, interest can be described as the outcome of the interaction between a person and a specific content (Krapp, 2000). Researchers further distinguish between situational and individual interest. The former refers to the attention and affective reaction in the situation triggered by environmental stimuli (Hidi, 1990). Individual interest describes a person's enduring predisposition to reengage with a specific content over time (Krapp & Fink, 1992). It can be conceptualized as a disposition or a current state (Krapp et al., 1992). Dispositional interests are enduring and general orientations, whereas current interests are concrete and based on interactions between internal and external conditions (Hidi & Baird, 1986).

Interest influences several learning-related aspects. Examples are attention (e.g., Ainley et al., 2002; Hidi, 1995; Hidi et al., 2004), goals (e.g., Harackiewicz et al., 2000; Pintrich & Zusho, 2002; Sansone & Smith, 2000) and the level of learning (e.g., Alexander & Murphy, 1998; Harackiewicz et al., 2002; Renninger & Hidi, 2002). There is also empirical evidence for the relationship between interest and prior knowledge (e.g., Baldwin et al., 1985; Tobias, 1994).

2.2.2.2 Self-Concept. *Self-concept* is a multidimensional and hierarchical construct with one global self-concept and several subordinate specific facets (e.g., Marsh et al., 1988; Shavelson et al., 1976). Self-concept is characterized by its evaluative character. There are investigations on the distinction between self-concept and self-esteem. However, recent research often treats these two global concepts as synonymous (Marsh & Martin, 2011). A positive self-concept is a desirable outcome in several psychological disciplines as well as education (Marsh, 2014; Marsh & Craven, 2006). In the latter domain, self-concept affects academic behaviors, choices, and achievement in and beyond school (Marsh, 2007, 2014; Marsh et al., 2006).

One specific facet of self-concept in education is the academic self-concept (e.g., Byrne, 1996). This self-perception is formed through experiences in educational environments (Guay et al., 2010; Craven & Marsh, 1997). Reciprocal effects between academic self-concept and achievement have been proposed by Marsh and Yeung (1998) and empirically tested (e.g., Guay et al., 2003; Marsh et al., 2005).

2.2.2.3 Self-Efficacy. *Self-efficacy* describes the judgment of one's belief of personal efficacy in the capability of executing cognitive, social, and behavioral actions and skills. (Bandura, 1977, 1986a, 1986b, 1997). Therefore, self-efficacy is related to specific situations and tasks. In contrast to self-concept, self-efficacy solely includes cognitive components (Peiffer et al., 2020). Bandura (1995) purported four sources of self-efficacy: earlier performance accomplishments, vicarious experience through observation of others, verbal persuasions, and evaluation of physiological and emotional information. Earlier performance accomplishments are the most essential sources (Bandura, 1981), but there is empirical evidence for the unique contributions of each of the four (e.g., Matsui et al., 1990). Moreover, external factors, for example, expectations of others (Bandura, 1986a), as well as internal and external cues (Gist & Mitchell, 1992), impinge one's self-efficacy. There is meta-analytic evidence of the importance of self-efficacy, specifically academic self-efficacy, for academic performance (r = .33, Honicke & Broadbent, 2016). It influences several aspects of learner behavior in educational settings, for example, the choice of activities and the level of performance or engagement (Linnenbrink & Pintrich, 2003; Schunk, 1985; Schunk & Pajares, 2002).

Academic self-efficacy primarily indicates one's perceived confidence to successfully perform a particular academic task (Bong & Skaalvik, 2003; Ferla et al., 2009). Besides the four sources of general self-efficacy, goal setting and attributional feedback play a role in connection with academic self-efficacy (Hodges, 2008). Each of the sources mentioned above is present in classrooms making the educational setting per se an essential factor for self-efficacy (Schunk, 1985).

2.2.2.4 Intrinsic and Extrinsic Motivation. People vary not only in their level of motivation but also in the orientation of that motivation (Ryan & Deci, 2000a). Based on the Self-Determination Theory (Deci & Ryan, 1985), it is possible to differentiate motivation into *intrinsic* and *extrinsic*.

Intrinsic motivation includes doing something because it is inherently interesting (Locke & Schattke, 2019). It is defined differently depending on the underlying theoretical model: Whereas the operant theory points to the reward in the activity itself (Skinner, 1953), learning theory and Self-Determination Theory focus on satisfying innate psychological needs (Hull, 1943; Ryan & Deci, 2000a). Meta-analytical evidence shows that intrinsic motivation plays a significant role in school achievement (Taylor et al., 2014).

Extrinsic motivation refers to doing something because of a separable outcome (Ryan & Deci, 2000a). Extrinsic motivation contrasts with intrinsic motivation by focusing on the instrumental value of the behavior. It can be defined as a means-end relationship in which doing something results in future value outside the task (Locke & Schattke, 2019).

A lot of educational activities are not inherently intrinsically interesting. Additionally, offering extrinsic rewards, for example, school grades, is proposed to inhibit students' will to learn intrinsically motivated (Covington & Müeller, 2001; Kruglanski, 1978). However, trying to encourage intrinsic values may also be discouraging. This is called the overjustification effect (Lepper et al., 1973). Students are more likely to value their task if there are task-oriented reasons for learning and when they are interested in the content (Covington, 2000).

Intrinsic and extrinsic motivation are differently connected to academic achievement. Lin and colleagues (2003) found that a high level of intrinsic motivation was positively related to grades for college students and that a moderate level of extrinsic motivation was better than a high level. This implied the possibility of a curvilinear relationship (Lin et al., 2003). There is similar evidence for elementary school students: Intrinsic motivation was associated with greater achievement, whereas extrinsic motivation showed a negative relationship (Lemos & Veríssimo, 2014). According to the Self-Determination Theory, internalizing and integrating values and behavioral regulations, the processes of taking them in and transforming them into one's own, answer how to motivate students to value activities (Deci & Ryan, 1985). This process leads to better persistence and self-perceptions (Ryan & Deci, 2000a).

By acknowledging motivational factors in learners and their relation to knowledge acquisition, more insight can be gained from mediating or moderating processes. Learners cannot be examined apart from their motivation concerning the learning process.

2.3 Knowledge Transfer

Knowledge transfer, or short *transfer*, is the ability to apply knowledge to new situations and contexts (Schwartz et al., 2008). Knowledge transfer enables learners to use the knowledge acquired in a specific context in other contexts to solve similar or dissimilar problems. This transfer includes transmitting knowledge to new content areas and physical, temporal, or functional contexts (Barnett & Ceci, 2002). It can also be described as the impact of learning in one context on learning and performance in other contexts (Perkins & Salomon, 1992) or the result of how learning in one task or situation influences the response in another task or situation (Adams, 1987).

Thus, the more learners transfer their knowledge, the broader the range of new problems they can solve in different contexts. This is an essential factor at school as well as in professional trainings. Accordingly, fostering transfer is among the central goals of education.

There are several fundamental distinctions (Chapter 2.3.1) and theoretical viewpoints (Chapter 2.3.2) of knowledge transfer. Additionally, a theoretical model of the impact of cognitive engagement in trainings on knowledge transfer is presented in Chapter 2.3.3.

2.3.1 Differentiations of Knowledge Transfer

Royer (1979) proposed at least four possible classifications of knowledge transfer that each contain two categories: First, lateral transfer includes a generalization of knowledge over a broad set of situations, whereas, in vertical transfer, the learned skill contributes only to a superordinate skill. Second, specific transfer with an apparent similarity between the learning and application situation can be differentiated from nonspecific transfer, in which no obvious stimulus elements are shared between both situations. The third distinction discerns literal transfer from figural transfer. The former describes the transfer of an intact skill, whereas the skill does not have to be intact for the latter. Last, the differentiation between near and far transfer depends on the similarity between the learning and the application situation.

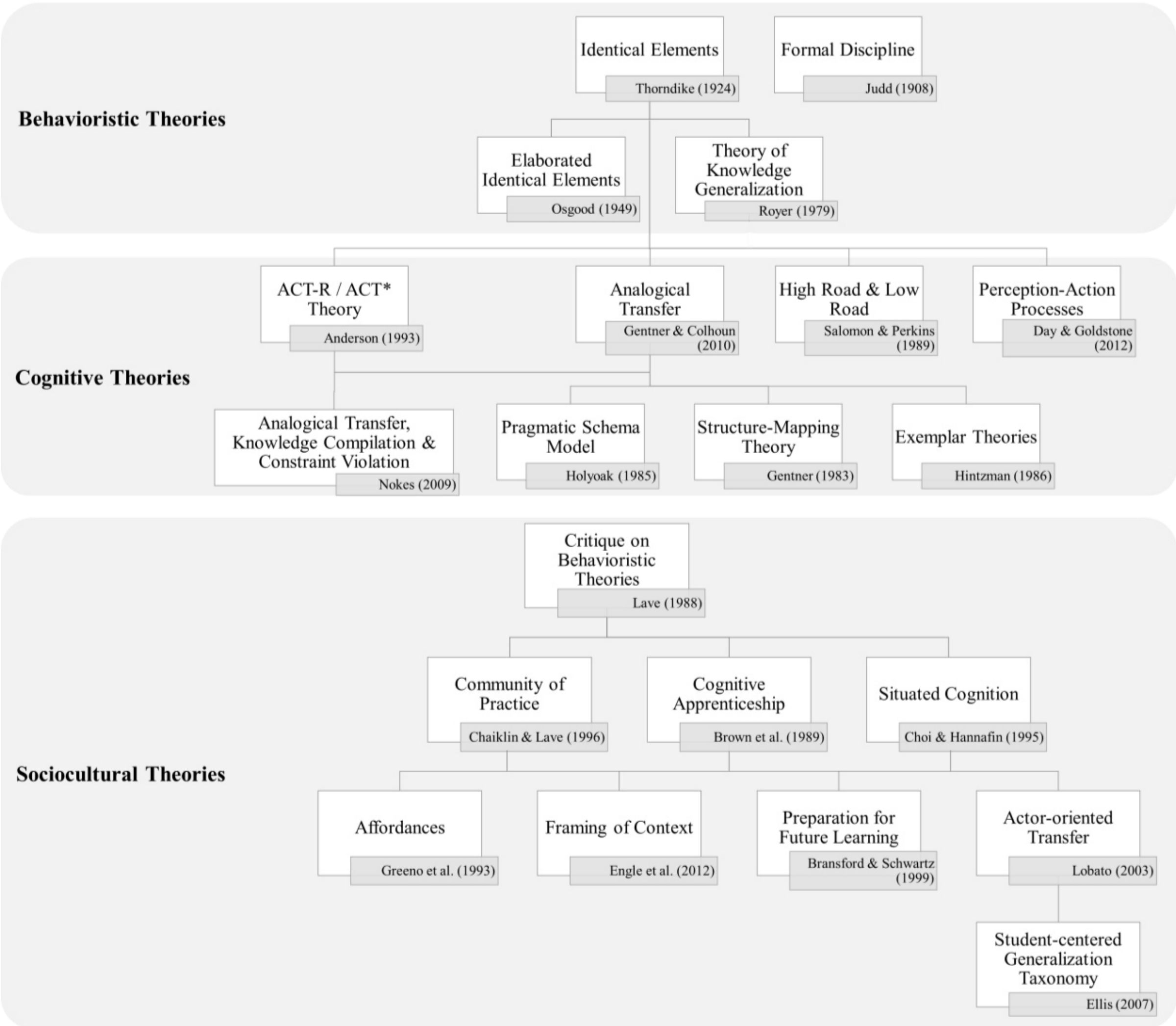
This distinction between near and far transfer is further explained in the following: Near transfer describes cases when knowledge is transferred to a mainly similar context. Empirical studies often found this type of transfer (Perkins & Salomon, 1992). Far transfer, in which knowledge is transferred to a dissimilar context, is rarely discovered (Haskell, 2001). The Taxonomy for Far Transfer (Barnett & Ceci, 2002) offers the possibility to distinguish these two types of transfer depending on the similarity of the learning and application contexts. It provides two overall transfer factors: the content, including what is transferred, and the context, meaning where it is transferred to (Barnett & Ceci, 2002). This model further differentiates six context factors (Barnett & Ceci, 2002): *Knowledge Domain*, *Physical Context*, *Temporal Context*, *Functional Context*, *Social Context*, and *Modality*.

2.3.2 Theoretical Views on Knowledge Transfer

Knowledge transfer has a long history of research. Theories on transfer are very diverse in their foci and approaches. However, they can be distinguished into three broad schools of thought (see Figure 3): These are behavioristic theories (Chapter 2.3.2.1), cognitive theories (Chapter 2.3.2.2), and sociocultural theories (Chapter 2.3.2.3). The three schools of thought and exemplary theories belonging to them are presented in the following.

Figure 3

Overview of Transfer Theories



Note. Overview of transfer theories, each with a reference example (own representation).

2.3.2.1 Behavioristic Theories. Traditionally, the discussion about knowledge transfer starts with the conflict between Thorndike and Judd and their rival theories. On the one side, Thorndike (1924; Woodworth & Thorndike, 1901) postulated that the crucial cognitive process in transferring knowledge is identifying identical elements in the source and the target domain. This theory is called the *identical elements theory*. After identifying such identical elements, knowledge acquired in one domain can be applied in a second domain (Thorndike, 1924). Identical elements are stimulus features that both tasks and situations share (Royer, 1979) or environmental as well as mental events and objects that differ in any respect between those situations (Butterfield & Nelson, 1989). Therefore, transfer is more likely between two similar tasks (Marton, 2006).

The theory of identical elements was further elaborated by Osgood (1949) and Ellis (1965), who described boundary conditions of most situations, resulting in restricting the theory to mostly near transfer. For example, Osgood (1949) separated the similarity of responses from stimuli similarities, resolving the paradox of identical elements. He formulated a three-dimensional theoretical surface (Butterfield & Nelson, 1989). According to this surface, positive transfer depends on the similarity of stimuli and responses, whereas there is no transfer when the stimuli are unrelated. Furthermore, negative transfer is described as similar stimuli but not responses (Butterfield & Nelson, 1989). The *theory of knowledge generalization* describes possible extensions to the theory of identical elements regarding far transfer (Royer, 1979). It postulates the possibility of far transfer when a class of problems is defined to which a particular skill can be applied and when it is possible to isolate such a set of defining features for this class (Royer, 1979).

On the other side, Judd (1908) purported that transfer is not only a function of similar features of two situations but also of how the learner deals with the learning situation (Marton, 2006). Transfer occurs due to teaching general rules in a domain, which then generalize to other domains (Lehman & Nisbett, 1990). In cases where a guiding principle is present, transfer is dependent on the extent a learner is accustomed to this principle (Brown & Kane, 1988). According to this, learning in formal demanding domains should enhance the quality of thinking in general (Inglis & Attridge, 2016). This theoretical view, initially developed by Plato, is called the theory of *formal discipline* (Lehman & Nisbett, 1990). There is empirical evidence for this theory (e.g., Fong et al., 1986; Lehman et al., 1988; Nisbett et al., 1987). However, most of the current transfer theories do not support the theory of formal discipline.

Up to this point, behavioristic theories dominated the research field of knowledge transfer. The most critical step in these theories is recognizing that one situation shares common environmental elements with another (Royer, 1979). They are limited to situations with clear and known relations between the initial stimuli and the transfer tasks (Kintsch, 1970; Royer, 1979).

2.3.2.3 Cognitive Theories. Cognitive theories concentrate on information processing in learning (Royer, 1979). These theories see the learner's memory as a highly structured system in which information is stored and retrieved systematically. Therefore, comprehension of contents is necessary, but not sufficient, for transfer (Royer, 1979). The likelihood of transfer is determined by the probability of retrieving relevant prior knowledge during the search process (Royer, 1979).

During the cognitive shift in psychological research, Singley and Anderson (1989) further enhanced the theory of identical elements. They formulated the *ACT* Theory* (Anderson, 1983a, 1987a) and the *ACT-R Theory* (Anderson, 1993). In these theories, the focus lies on cognitive skills that are composed of production rules in a system. This production system has several critical features (Anderson, 1993): Production rules are abstract modular pieces of knowledge characterized by their condition-action-asymmetry and strung together by the learner's goal-setting. These production rules can be used as a "*modified elements theory* *of transfer*" (Anderson & Singley, 1993, p. 184). Instead of stimulus-response bonds between the learning and application situation, transfer is described as a function of how the learner represents knowledge in the form of knowledge chunks and productions (Anderson & Singley, 1993). This makes transfer more likely when the same production rules apply in the learning as in the transfer context.

Salomon and Perkins (1989) distinguished two distinct but related knowledge transfer mechanisms. *High-road transfer* includes deliberate reflective processing and mindful abstractions from the initial learning context (Perkins & Salomon, 1992). It consists of an intentional search for connections. *Low-road transfer* depends on pattern recognition and the reflexive triggering of semi-automatic processes (Perkins & Salomon, 2012). This mechanism occurs when the conditions of the application setting are similar to those in the learning situation (Perkins & Salomon, 1992). Three processes can support both roads (Perkins & Salomon, 2012): detecting, electing, and connecting. The process of *detecting* discerns the possibility of a connection between the learning and application situation that the learner needs to *elect* to pursue further. *Connecting* describes the process of identifying relevant relations between the two situations. This is supported by high surface similarity (Day & Goldstone, 2012). Failures in transfer are explained by missing an adequately deep understanding of the contents (Chi & VanLehn, 2012). Furthermore, another prerequisite for these processes is the successful reconstruction of schemas of the source (Perkins & Salomon, 2012).

Analogical transfer is based on the importance of analogical problem-solving (Gysin & Brovelli, 2021). It is premised on transmitting knowledge features from one problem to another with similar structural links by drawing on previous encounters (Gary et al., 2012). Analogies include matching relevant parallels between the problem and the situational structures by the learner (Blanchette & Dunbar, 2000; Gentner et al., 2001). They can be drawn between and within knowledge domains. Learning to solve one problem enhances the solving process of another problem depending on the similarities of the learner's mental representations of the two problems (Marton, 2006). Transfer due to analogies between the learning and application situation has been investigated through well-defined puzzle problems that are either homomorphic or isomorphic (Hayes & Simon, 1977; Reed et al., 1974). There are three main classes of theories on analogical transfer (Reeves & Weisberg, 1994). These are the *pragmatic schema model* (Holland et al., 1986; Holyoak, 1985; Holyoak & Thagard, 1989), the *structure-mapping theory* (Gentner, 1983, 1989; Gentner & Gentner, 1983), and *exemplar theories* (e.g., multiple-trace model, Hintzman, 1986, 1988). Research on analogical thinking grounded on the pragmatic schema theory in ill-defined complex problem-solving tasks. This was done by Gick and Holyoak (1980) with the radiation problem. It led to the conclusion that problem-solving is possible based on one analog. However, developing an adequate schema was most important for successful transfer (Reeves & Weisberg, 1994). Analogical transfer, predicated on the structure-mapping theory, includes several processes: retrieval, mapping, evaluation, abstraction, and re-representation (Gentner & Colhoun, 2010). Calling an analogous case from memory based on the confrontation with a new case is called retrieval (Keane et al., 1994). The most central process, mapping (Gentner, 1983; Gentner & Colhoun, 2010), entails the structural alignment between two cases and their implications. During this process, aspects that align are searched, clustered, and mapped when found (Gentner & Colhoun, 2010; Keane et al., 1994). The result is a set of correspondences between matches of both cases. The found analogy is evaluated (e.g., Forbus et al., 1997; Holyoak & Thagard, 1989; Keane, 1996) and often abstracted from the cases' structures. In the process of re-representation, one or both cases are adapted to better fit the analogy (Gentner & Colhoun, 2010; Gentner et al., 2003). Last, the third class of theories discerns from the other by focusing on concrete problems and experiences (Reeves & Weisberg, 1994). One example of such exemplar theories is the multiple-trace model (Hintzman, 1986). This model presumes that each event includes its own memory trace. Therefore, repetition of previously learned content does not strengthen the prior representation but produces a new coexisting trace (Hintzman, 1976, 1986; Hintzman & Block, 1971). Transfer enhances by providing prototypes of old exemplars (Hintzman, 1986; Rosch et al., 1976) and increasing the number of training exemplars (Homa et al., 1973; Homa et al., 1981). For this, learners encode the learning task's surface, structural details, and contextual information (Reeves & Weisberg, 1994).

Learners can also compare two cases that are neither well understood. This technique is called analogical encoding (Gentner et al., 2003) or analogical bootstrapping (Kurtz et al., 2001). This comparison helps to understand the underlying structure of both cases (Loewenstein et al., 1999) by drawing attention to common principles and schemas (Gick & Holyoak, 1983). Promoting the abstraction of existing schemas facilitates solving problems in different contexts (Gentner et al., 2003). Likewise, such between-domain analogies can be based on identifying salient similarity, the similarity of the learner's underlying representations (Vosniadou & Ortony, 1989).

Focusing on the shift from abstract and high-level conceptualization processes to *perception-action processes*, Day & Goldstone (2012) provided an overview of the connection between perceptual learning processes and knowledge representation. In traditional cognitive theories, analogical encoding is an active and deliberate process. However, perceptions are fast and automatic and include a combination of perceptual and conceptual content (Day & Goldstone, 2012). They are influenced by prior knowledge and show parallels to aspects of procedural knowledge (Gauthier et al., 2010; Shiffrin & Schneider, 1977). These perceptions are stored as mental models that are simplified spatial and mechanical representations of situations. They allow reasoning about relations between situations (e.g., Dreistadt, 1969; Johnson-Laird, 2006). These perceptual mental representations are essential in transferring cases that appear dissimilar on the surface but are represented similarly (Day & Goldstone, 2011, 2012). Therefore, perceptions provide a basis for the linking of situations. These effects are not restricted to visual representations. Indeed, the relationship between physical actions and thought is incorporated (e.g., embodied cognition, Barsalou, 1999; Goldin-Meadow, 2005; Wilson, 2002). The authors also purported the importance of prior knowledge in recognizing and taking advantage of the deep structural content in cases where the learner does not recognize the deep structure (Day & Goldstone, 2012).

Building on the beforementioned cognitive theories, Nokes (2009) presented three profiles of processes that can be triggered in knowledge transfer. They depend on the type and the representation of the knowledge to be transferred. The processes are *analogical transfer*, *knowledge compilation*, and *constraint violation*. Analogical transfer, as explained above, includes the retrieval of prior exemplars and the creation of mappings between these and the current problem (Catrambone, 2002; Holyoak & Koh, 1987). It is facilitated by surface and structural similarities (Catrambone & Holyoak, 1989; Gentner & Gentner, 1983; Gick & Holyoak, 1983). The inferences drawn by this process are typically assumed to be declarative representations (Nokes, 2009). Novices apply this process to near-transfer problems that look similar on the surface. The second process, knowledge compilation, entails translating prior declarative knowledge into procedures to solve new problems (Nokes, 2009). The generation of production rules includes step-by-step interpretations of declarative knowledge representations (Anderson, 1982, 1983b, 1987b). These rules are widely applicable but require a complicated and lengthy application process (Taatgen & Anderson, 2002). Therefore, the process is only triggered when no accessible exemplar knowledge or extensive adaptation is required. Third, constraint violation contains different sets of cognitive processes to realize knowledge transfer from declarative to procedural knowledge (Ohlsson, 1996; Ohlsson et al., 1992). Constraint violation consists of the repeated processes generate, evaluate, and revise (Nokes, 2009). The learner uses prior knowledge of domain constraints to evaluate and correct task performances. Transfer is the process in which the learner uses prior constraint knowledge to identify errors generated in the performance of a new task. This process requires multiple iterations of the repeated processes and is only triggered when the learner has no accessible exemplars and tactical knowledge (Nokes, 2009).

Behavioristic, as well as cognitive theories, have both strengths and weaknesses. Whereas behavioristic theories are better suited in their predictive specificity, their structure is relatively rigid. They provide guidelines for developing and sequencing instructional events but not for transfer from school-learned material to real-world events (Royer, 1979). Cognitive theories are broader and yield explanations for knowledge transfer in various situations for both near and far transfer.

2.3.2.4 Sociocultural Theories. In the late 1980s, critique started on the classic behavioristic and cognitive transfer theories (Marton, 2006). This critique mainly targeted the researcher's role in defining learned content and transfer tasks (Lave, 1988). For example, the traditional two-problem transfer experiment disregards other possible relations to situations (Marton, 2006). Up to this point, research on knowledge transfer was based on a functionalist view with the image of knowledge as tools stored in memory (Marton, 2006). However, later theories focused more on the situational aspects of knowledge and transfer (Day & Goldstone, 2012; Roorda et al., 2015). Knowledge is seen as situated and dependent on the context (Boaler, 2002; Lave, 1988). Sociocultural theories, including the *community of practice* (Chaiklin & Lave, 1993; Frade et al., 2009; Lave, 1988; Wenger, 1998; Wenger et al., 2002), *cognitive apprenticeship* (Brown et al., 1989; Greeno et al., 1993), and *situated cognition* (Choi & Hannafin, 1995; Hennessy, 1993; Suchman, 1987) developed from Lave's (1988) initial critique.

Knowledge transfer, also framed as the generality of knowing (Greeno, 1997), is based on one's ability to participate in activities in different situations that provide *affordances*

(Greeno et al., 1993). Affordances (Gibson, 1977) are features of the environment that make particular activities, for example, knowledge transfer, possible (Day & Goldstone, 2012; Marton, 2006). Affordances are non-dualistic; this means they presuppose both an actor and an acted upon (Marton, 2006). Greeno et al. (1993) presented a view of transfer as a conjunction between affordances in the learning and the transfer situation. Transfer is constituted by the person in the situation and the situation for the person (Greeno, 1997). Therefore, learning and transfer are situated and depend on the social context.

Adding to the situational dependency of knowledge and its transfer, Engle and colleagues (2012) purported the importance of *framing the social context*. Engle (2006) argued that learners need not only to possess relevant knowledge but also to choose to use their knowledge to transfer. This choice can be influenced by social framing (Engle, 2006). To enable transfer, an expansive social framing of the context is needed to create intercontextuality between the learning and transfer contexts (Bloome et al., 2005; Engle, 2006; Leander, 2001). This is possible by connecting prior and future settings and emphasizing the student's authorship of their knowledge (Engle, 2006; Engle et al., 2012; Engle et al., 2011; Hammer et al., 2005).

Broadening the theoretical conceptualization of transfer, the learner's *preparation for future learning* was included as an essential aspect of knowledge transfer (Bransford & Schwartz, 1999). It incorporates assessing the ability to learn successfully in new environments (Schwartz et al., 2005; Schwartz & Martin, 2005). Furthermore, it answers how knowledge acquired in one situation transfers in and enables learners to learn from new unknown resources (Belenky & Nokes-Malach, 2012; Nokes-Malach & Mestre, 2013). This focus allows researchers to notice evidence of positive transfer that might otherwise be hidden in the traditional researcher-oriented view of knowledge transfer. Bransford and Schwartz (1999) recommended assessing transfer by directly exploring learners' abilities to learn new information while connecting them to prior learning. This focus on transfer has also been discussed before but not this explicitly (e.g., Bereiter, 1990; Greeno et al., 1993). The assessment of transfer regarding the preparation for future learning demands specific experimental designs. One way is the double-transfer paradigm (Belenky & Nokes-Malach, 2012; Stratton, 2020). In this paradigm, students receive one of two instructions, and half of these students then receive additional learning resources afterward, often operationalized as inventing versus tell-and-practice (Schwartz & Martin, 2005). Following, all of them complete a transfer problem. The preparation for future learning approach can be closely connected with motivational factors. For example, by looking at achievement goals, the mastery-approach orientation is connected with better preparation for future learning (Belenky & Nokes-Malach, 2012).

Another sociocultural concept of knowledge transfer is the *actor-oriented transfer* (Lobato 2003, 2008, 2012). It also criticizes the traditional two-problem transfer paradigm (Lobato, 2003). In this, the researchers, as experts, look for improved performance between learning and transfer tasks (Reed, 2012). Hence, classical studies include a privileged perspective of observers and rely on models of expert performance (Lobato, 2006). Traditional transfer paradigms do not account for structuring sociocultural environments or material artifacts (e.g., Guberman & Greenfield, 1991; Lobato, 2012). However, from the actor-oriented transfer point of view, knowledge transfer is inspected from the learner's perspective (Lobato et al., 2012). Researchers look for the impacts of prior activities on the current activity and how actors construe situations as similar. Therefore, knowledge transfer is the generalization of learning or the effect of learners' prior activities on other activities in novel situations (Lobato, 2003). In this theory, knowledge transfer is typically examined qualitatively with ethnographic methods by scrutinizing given activities to indicate influences of prior activities (Lobato, 2006; Lobato & Siebert, 2002).

A further advancement of the actor-oriented transfer was made by combining this concept with a taxonomy for generalization (Ellis, 2007a). Generalization has previously been described as identifying commonalities (Dreyfus, 1991; Ellis, 2007b) and shares features with the classical definition of lateral transfer (e.g., Anderson et al., 1995; Ellis, 2007b). Ellis (2007a) identified cognitive actions during generalizing (Ellis, 2011): The resulting *student-centered generalization taxonomy* extends the actor-oriented transfer perspective to different types of generalizations. These are generalizing actions and reflection generalizations (Ellis, 2007a). Generalizing actions are mental acts based on learners' activities. Reflection generalizations are public statements explicitly stating common properties or relations. They are linked to the learners' generalizing actions (Ellis, 2007a).

The beforementioned theoretical views on knowledge transfer show a variety of foci, experimental designs as well as measurements of transfer. Three main schools of thought arise from behavioristic, cognitive, and sociocultural theories. However, it becomes clear that there are differences between and within the three schools of thought. This makes it difficult to determine an overall definition of knowledge transfer that can be used in research.

2.3.3 Interventions and Trainings to Foster Transfer

Similar to the concepts presented so far, various possible definitions of *intervention* or *training* exist. In this dissertation, we concentrate on pedagogical trainings with three key characteristics (Fries & Souvignier, 2009): First, there must be some self-managed or guided repetitive practice. Second, the content of the training includes the acquisition of knowledge to enhance knowledge or skills. The last feature covers the structure. The organized and temporal limited knowledge acquisition phase characterizes a training. In this context, a training is a structured and temporal limited intervention in which knowledge is acquired through practice (Fries & Souvignier, 2009).

There are several frameworks on the implication of instructional methods on knowledge acquisition. However, models that also include knowledge transfer are scarce. A recent model, the Interactive-Constructive-Active-Passive framework (ICAP framework, Chi, 2009; Chi & Wylie, 2014), includes learners' cognitive engagement activities, operationalized as overt student behavior, in answer to instructional methods. This behavior is divided into four hierarchical modes of cognitive engagement. Chi and Wylie (2014) proposed that the first mode describes an overtly passive learning situation without students showing any behavior related to their learning (Henderson, 2019). The expected cognitive outcome of this passive mode is the recall of learned content without application. The active mode, as the second category, contains overt motoric actions by students that are confined to the provided information, for example, pointing to what they are reading or underlining a text (Alibali & DiRusso, 1999; Chi & Wylie, 2014; Katayama et al., 2005). This mode enables students to apply their knowledge to non-identical but similar contexts. Third, generating additional externalized outputs, e.g., asking questions (Graesser & Person, 1994) or taking notes (Trafton & Trickett, 2001), are part of the constructive mode (Henderson, 2019). As a result, learners can transfer knowledge to novel contexts or distant problems. The highest mode of the hierarchy is the interactive mode. It contains working constructively in dialogues with partners, including defending a position (Schwarz et al., 2000), asking questions (Webb, 1989), or correcting others (Hogan et al., 1999). The cognitive outcome of this mode is the co-creation of new ideas that go beyond the capacities of one sole learner. Knowledge transfer can be anticipated as a cognitive outcome in the active, constructive, and interactive mode.

The four modes of cognitive engagement in the ICAP framework are independent of the used instructions themselves. This means various engagement activities can occur while learning, regardless of the applied methods (Chi & Wylie, 2014). Nevertheless, instructional methods can be assigned to the four modes based on the expected learner's reaction. The ICAP framework allows the connection between instructional methods used in trainings and their cognitive outcomes in the form of knowledge transfer.

3 Methodological Approaches

In this chapter, I describe the methodological approaches used in this dissertation. First, the general scope of literature reviews is explained in Chapter 3.1. Following, the aspects and the process of a meta-analysis as an additional quantitative approach are described in Chapter 3.2. Information on the used models in the meta-analyses of this dissertation, the cross-lagged panel model and the mediation model, are outlined in Chapters 3.3 and 3.4, respectively.

3.1 Literature Review

There are several definitions of *literature reviews* (e.g., Siddaway et al., 2019). A connecting feature of all these definitions is that they use publications containing preliminary information to code and synthesize them (Cooper et al., 2018). There are several characteristics in which literature reviews can differ. Examples are the focus, the goal, and the perspective (Cooper et al., 2018). However, the main reason for a literature review is the same: They include an overall impression of the existing evidence for a research question (Siddaway et al., 2019).

The main distinctions can be made on two levels: the research process and the methodology. The first distinction is placed on the level of the research process, differentiating between a narrative review and a systematic review. A systematic review is characterized by providing explicit information on the search and inclusion process through documented search phrases and inclusion as well as exclusion criteria (Siddaway et al., 2019). With this information, the systematic review is reproducible and can be replicated. Narrative reviews, on the other hand, profit from a broader focus and more comprehensive coverage (Collins & Fauser, 2005). The second distinction regards the methodological approach of synthesizing data. It is possible to conduct a review solely qualitatively or include quantitative measures. Both the systematic and the narrative reviews are qualitative literature reviews.

3.2 Meta-Analysis

A quantitative review, also called *meta-analysis*, is indicated when the review aims to synthesize studies that empirically test the same hypotheses. In this case, information about empirical evidence, the effect sizes, is synthesized and statistically analyzed (Borenstein et al., 2009; Schmidt & Hunter, 2015).

The purposes of a meta-analysis are diverse (Bulpitt, 1988; Fagard et al., 1996; Hedges, 1992; Sacks et al., 1987; Thompson & Pocock, 1991). One example is the increase of statistical power based on the increase in observations. Furthermore, included effects are weighed by their precision (Cheung & Vijayakumar, 2016). This enables researchers to examine variabilities between studies, resolve uncertainties, and perform subgroup analyses.

There are different statistical models of meta-analyses (Cheung, 2015b). For example, it is possible to differentiate between the *Fixed-Effects-Model* and the *Random-Effects-Model*. The Fixed-Effects-Model assumes that the effects from the included primary studies come from the same homogeneous population (Schwarzer et al., 2015). The Random-Effects-Model, on the other hand, does not involve this restriction. Therefore, it additionally accounts for differences between included studies (Cooper et al., 2018). Moreover, models can encompass statistically independent or dependent effect sizes in their primary studies. Dependent effect sizes result from reporting several measurements or measurement points of the same population. Including such effect sizes violates a central assumption of meta-analyses (Cheung, 2019; Cheung & Chan, 2005). Several possible solutions exist, such as structural equation models (e.g., Cheung, 2015b) and robust variance estimation (e.g., Hedges et al., 2010; Tanner-Smith et al., 2016). To compute a main effect over studies, there is a need for corresponding effect sizes that include the defined relationship of interest (Cooper, 2015). It is crucial only to examine effect sizes that represent the same relation. Otherwise, the results will not be trustworthy.

3.2.1 Conducting a Meta-Analysis

The conduction of a meta-analysis follows a pre-set procedure from the systematic search in databases to the statistical analysis. One can follow several guidelines in the process (e.g., PRISMA Reporting Guidelines; Arya et al., 2021; Page et al., 2021). The main steps in the process of a meta-analysis, after the problem or research questions are stated, are the systematic and exploratory search, the screening of titles, abstracts, and full texts, the coding of information, and the statistical analysis (Cooper, 2015).

The steps of a meta-analysis are outlined in the following: The first step in conducting a meta-analysis, after formulating a problem statement, is the systematic search in databases and the exploratory search in additional sources. The main goal of this step is the coverage of "all the research conducted on the topic of interest." (Cooper et al., 2018, p. 12). A combination of search words is formed and used in several databases to conduct the systematic search. For the screening procedure, inclusion criteria are formulated and documented. These criteria specify the aspects the primary studies need to contain to be included in the meta-analysis (Field & Gillett, 2010). Such aspects are the empirical relation, the design and methods, the features of the sample, and the required statistical data to compute the effect size. After inclusion, the systematic coding of information takes place. This coding is based on a coding protocol with a clear delineation of the critical information. Information entails the identification of the included report, the setting and the sample, the methodology and dependent measures, and the effect size information. Last, statistical analysis is applied (Cheung & Vijayakumar, 2016). This analysis depends on the predefined model, the effect sizes, and the main and moderator hypotheses (Field & Gillett, 2010).

3.2.2 Advantages and Disadvantages of Meta-Analyses

There are several advantages based on the method of meta-analyses. The information gained is spanned over a broader range by integrating primary studies. Therefore, the results of meta-analyses provide better estimates of the relations in the population than single studies (Schmidt & Hunter, 2015). In addition, the precision and accuracy of estimates can be improved. This is premised on the increased data used in a meta-analysis that provides more statistical power to detect effects than separate independent studies (Schmidt & Hunter, 2015). Meta-analyses also help researchers resolve inconsistencies in research findings and enable them to identify moderating or mediating variables that may explain the reasons for these discrepancies (Walker et al., 2008). Last, the systematic procedure ensures replicability (Aytug et al., 2012; Cheung & Vijayakumar, 2016).

However, the results of meta-analyses must be interpreted with some caution regarding several common problems. First, the conduction can be problematic if the selection of studies is biased by not identifying all possible studies (Greco et al., 2013). Further, the validity and quality of the included studies influence the results of the meta-analysis. This effect is called *garbage-in-garbage-out* (Sackett et al., 1986). The next problem is the *publication bias* or *file drawer problem* (e.g., Walker et al., 2008). This means published records are more likely to include significant or confirming results. Non-significant or contradictory results are published less often (e.g., Easterbrook et al., 1991; Francis, 2013; Song et al., 2010). The journals' or authors' prestige might also affect the inclusion of studies. Furthermore, primary studies often lack internal, construct, and external validity (Bobko & Stone-Romero, 1998). Studies with a small sample size bear an additional risk: They typically have low statistical power and large standard errors (Bobko & Stone-Romero, 1998). Last, it is possible that the heterogeneity between the methods used in primary studies and their data analyses cause misleading inferences about the relationship between variables (Greco et al., 2013; Murphy, 2017). As a result, a meta-analysis based on these studies can produce very heterogeneous effect sizes, and the findings may lead to erroneous inferences. Thus, methodologists argued that researchers should conduct sensitivity analyses to examine the heterogeneity in effect sizes of studies included in the analysis (Cooper et al., 2018).

3.2.3 Meta-analytic Structural Equation Models

Most meta-analyses investigate data via univariate models. Recently, the possibility of analyzing multivariate models in the form of structural equation models was developed. In the *Meta-analytic Structural Equation Model (MASEM)*, structural equation models (SEM) are fitted to meta-analytic data (Jak et al., 2021). With this, a MASEM combines the strengths of both meta-analyses and SEM (Cheung, 2015b). Typically, a MASEM consists of two stages (Cheung, 2015b): In the first stage, the correlation matrices from primary studies are all combined into one pooled correlation matrix. This matrix is then fitted by a SEM in the second stage (Jak & Cheung, 2020). In contrast, in a one-stage MASEM, the SEM is directly fitted to the correlations from primary studies. Therefore, it can be classified as an extension of the random-effects meta-analysis of correlation matrices in the first stage of the traditional two-stage MASEM (Cheung, 2014). Additionally, it enables testing moderator hypotheses with categorical and continuous moderators (Cheung, 2014).

Meta-analyses allow for the integration of several results from primary studies that provide a more profound insight into research questions. This leads to more trustworthy results (e.g., Schmidt & Hunter, 1996). However, problems can occur when meta-analyses are not conducted properly. This includes the use of inappropriate primary studies, methods, or approaches. Nevertheless, meta-analyses help provide answers with a profound data basis.

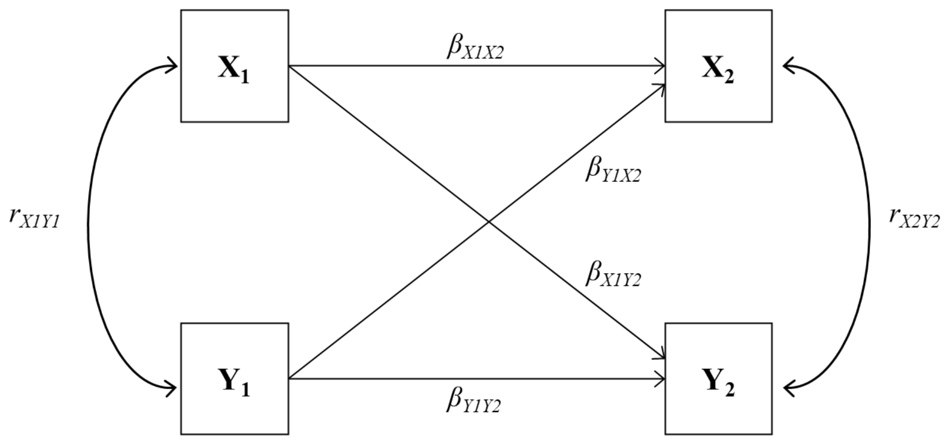
3.3 Cross-Lagged Panel Model

In several research areas, *cross-lagged panel models* (CLPM, Hamaker et al., 2015) are widely used because of the possibility of investigating cross-lagged effects while controlling for autoregressive effects of at least two variables (Kenny, 2005; Newsom, 2015; Rosel &

Plewis, 2008; Selig & Little, 2012). The CLPM provides information on autoregressive relations and time-lagged regressions of two variables (see Figure 4, Berry & Willoughby, 2017). Autoregressive relations contain information on rank-order stability (Mund & Nestler, 2019), and such models additionally control for the stability of the included constructs (Hamaker et al., 2015). On the other hand, time-lagged regressions can be interpreted as predictive relations over time between both variables (Berry & Willoughby, 2017). These are typically used as evidence of lead-lag or bidirectional relations (Berry & Willoughby, 2017).

Figure 4

Schematic Representation of a Cross-Lagged Panel Model



Note. Cross-lagged panel model with the cross-sectional (r_{C1P1} and r_{C2P}), auto-regressive (β_{C1C2} and β_{P1P2}), and cross-lagged paths between two variables at two measurement points (β_{C1P2} and β_{P1C2}).

Many studies in educational psychology use the CLPM approach for practical reasons (Usami et al., 2019). In contrast to other more advanced models, two waves of data are enough for a CLPM (Kenny, 2005). This makes it a helpful approach from a practical perspective (Hamaker et al., 2015). Additionally, from the statistical point of view, CLPMs are more suited to evaluate possible causal relations than simple cross-lagged correlations or multiple regressions (Hertzog & Nesselroade, 2003). This advantage is premised on the autoregressive effects of the two variables in a CLPM. The inclusion of these effects controls for their influence on the estimates of the cross-lagged paths (Orth et al., 2021).

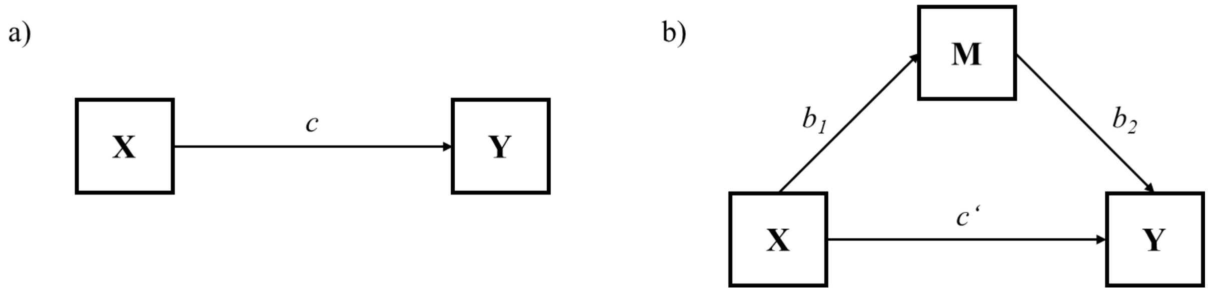
Nonetheless, the CLPM makes some assumptions that can be criticized in applied contexts (Mund & Nestler, 2019): The CLPM does not include any measurement-error variance for indicators (Lucas, 2023; Lüdtke & Robitzsch, 2022). It also assumes that all participants vary across one group mean without stable between-person differences (Mund & Nestler. 2019). This can be problematic because a CLPM does not distinguish within- and between-person variance (Berry & Willoughby, 2017; Lucas, 2023). Therefore, they may be conflated (Mund & Nestler, 2019). Additionally, the CLPM does not capture mean-level changes in the observed constructs over time and whether there are differences in this change between persons (Mund & Nestler, 2019; Usami et al., 2019).

3.4 Mediation Analysis

In its simplest form, *mediation* represents the addition of a third variable to the relation between two variables. In this relation, the independent variable causes the mediator, and the mediator causes the dependent variable (Igartua & Hayes, 2021). Therefore, "a mediator is a variable that is in a causal sequence between two variables" (MacKinnon et al., 2007, p. 595). Researchers testing for mediation paths usually make causal statements about these relationships (Huber, 2020; Iacobucci, 2008; Lee et al., 2019). Figure 5 displays a schematic representation of a simple mediation model. In such a model, three effects are essential (Agler & De Boeck, 2017): The *total effect c* of X on Y is displayed in Figure 5 a). This *total effect* is then partitioned into the *direct effect c'* of X on Y and the *indirect effect* transmitted through the mediator, paths b_1 and b_2, as seen in Figure 5 b) (Fritz et al., 2012).

Figure 5

Schematic Representations of a Simple Mediation Model



Note. a) model with a total effect *c* of X on Y. b) mediation model with direct effect *c'* and indirect effects b_1 and b_2. X – independent variable, Y – dependent variable, M – moderator variable.

There are several approaches to determine whether the mediator yields a significant effect (Agler & De Boeck, 2017; VanderWeele, 2016). The traditional approach was first introduced by Baron and Kenny (1986). They examined whether a direct effect is significantly reduced or becomes nonsignificant when subtracting the indirect effect from the former total effect in a model. This method has been criticized (e.g., Fritz & MacKinnon, 2007). The Sobel test (Sobel, 1982) is another method to test for mediation effects. This test is conducted by calculating the ratio of the point estimate to its standard error and determining the *p*-value under the null hypothesis of no indirect effect (Igartua & Hayes, 2021). It also has been criticized for its assumption of a normal sampling distribution (e.g., Hayes, 2009). Some alternatives have been proposed, including Bayesian methods (Yuan & MacKinnon, 2009) or the percentile bootstrap confidence interval using the PROCESS procedure (Hayes, 2022).

Researchers generally distinguish between complete and partial mediations. A significant path coefficient *c'* in combination with a significant indirect effect speaks for partial mediation (MacKinnon et al., 2007). Others purported up to five possible outcomes of mediation analyses (Zhao et al., 2010): The *complementary mediation* in which the indirect and direct effects point in the same direction. Second, if the direct and indirect effects point in opposite directions, it speaks for a *competitive mediation*. Third, *indirect-only mediation* entails only an indirect effect but no direct effect. Fourth, if only the direct effect exists, it is called a *direct-only nonmediation*. Last, the *no-effect nonmediation* includes neither direct nor indirect effects.

4 This Dissertation

This dissertation investigates three different aspects of knowledge. Starting with predictive relations between two knowledge types, this dissertation also covers mediation effects on learning outcomes and the impact of interventions on knowledge transfer. All these processes are part of the construct of knowledge but simultaneously encompass goals in educational settings that each are pursued individually. However, the beforementioned processes are challenging to examine in primary research because of their broadness. Therefore, in this dissertation, I investigated them through three separate meta-analyses.

Study 1 focuses on the predictive relations between prior and later conceptual and procedural knowledge in the knowledge domain of mathematics. This study used a one-stage MASEM as a cross-lagged panel model. We were interested in the path coefficients of the cross-lagged paths. According to the Iterative Model (Rittle-Johnson et al., 2001), we expected positive, bidirectional relations between both knowledge types. Several moderators, e.g., the mean age and school level, were investigated for their influences on the main effect.

Study 2 examines motivational mediators between prior knowledge and later knowledge after learning. We answer to what degree knowledge scores remain stable over time within groups of learners and whether motivational variables mediate this relation. It was hypothesized that prior knowledge as a predictor is positively associated with motivation. Motivation is expected to be a mediator variable explaining stable interindividual differences in knowledge scores after learning. Moderator variables were the level of specificity of motivation and the knowledge types, for example.

Study 3 presents a meta-analysis of the effectiveness of transfer interventions on knowledge transfer compared to control interventions. We investigated the advantage of instructional methods in transfer interventions for school students compared to control groups and groups that received fewer instructional methods. It was hypothesized that instructional transfer interventions foster transfer better than control interventions, with a medium-sized effect. We formed a theoretical matrix based on the learners' cognitive engagement (Chi & Wylie, 2014) and two cognitive transfer processes to assess moderator effects. The transfer distance based on the Taxonomy for Far Transfer (Barnett & Ceci, 2002) was also used as a potential moderator. Moreover, other exploratory moderators were investigated.

5 Study 1: The Predictive Relations Between Conceptual and Procedural Knowledge in Mathematics: A Research Synthesis Using Meta-Analytic Structural Equation Modelling

5.1 Abstract

Mathematical knowledge demonstrably is a strong predictor of success in several periods of life. Researchers in this domain often differentiate between conceptual and procedural knowledge. There are at least four theoretical views on the predictive relations between conceptual and procedural knowledge: The *Inactivation View*, the *Concepts-first Theories*, the *Procedures-first Theories*, and the *Iterative Model*. This meta-analysis analyzes the predictive relations between the two knowledge types to determine which of these theoretical views is the most appropriate in mathematics. Using the one-stage meta-analytic structural equation model approach, we meta-analyzed 51 independent correlation matrices from 34 studies (N = 7253) in a cross-lagged panel model. The included studies reported quantitative standardized effect sizes of the predictive relations at two or more measurement points in the knowledge domain of mathematics. Additionally, we conducted several moderator analyses to test the generalizability of our findings, e.g., concerning age and subdomains. We found significant and almost symmetrical predictive relations between conceptual and procedural knowledge (β_{C1P2} = .270, 95% CI [.221, .318], β_{P1C2} = .220, 95% CI [.172, .268]). This finding supports the Iterative Model. The results are stable across the examined moderators. There are several limitations; for example, the decision to use a cross-lagged panel model. Concluding, the findings of this meta-analysis offer insights into possible implications for mathematics education and future research.

5.2 Introduction

"The relationship between conceptual understanding and computational skill is one of the oldest concerns in the field of psychology of mathematics" (Resnick & Ford, 1981, p. 246). Over the last decades, the importance of the relation between conceptual and procedural knowledge has been highlighted across many strands of research and practical settings: Cognitive scientists have devised conceptual and procedural knowledge as two different knowledge types (e.g., Hiebert & Lefevre, 1986). Developmental psychologists investigated the naturally occurring developmental patterns of conceptual and procedural knowledge (e.g., Karmiloff-Smith, 1992). Educational psychologists analyzed how conceptual and procedural knowledge can be taught to facilitate further learning (e.g., Klahr et al., 2007; Rittle-Johnson et al., 2015). In the practical field of education, enhancing conceptual and procedural knowledge is a central learning goal throughout educational levels and domains. So, current guidelines for effective instruction in science education recommend teaching both knowledge types (National Research Council, 2012; Zoupidis et al., 2016).

Despite the great effort to understand conceptual and procedural knowledge development during the past years, the longitudinal associations are still under debate. The general predictive relation between domain-specific prior knowledge and knowledge after learning has been meta-analytically tested (Simonsmeier et al., 2022). According to Simonsmeier and colleagues (2021), there is evidence for the predictive relation between prior knowledge and later knowledge after learning. However, up to this point, it lacks meta-analytical evidence for the differentiation in conceptual and procedural knowledge in this relation. Therefore, this meta-analysis answers how strong the predictive relations between conceptual and procedural knowledge are. We aim to fill this gap by summarizing existing evidence with meta-analytic structural equation modeling.

Generally, it has been shown that mathematical knowledge strongly predicts achievement in various areas. For example, mathematical knowledge in preschool is associated with later achievement in mathematics and other domains (Claessens & Engel, 2013). Further, mathematical knowledge in early childhood is positively connected with socio-economic status and the evaluation of risks or benefits associated with choice options in adulthood (Reyna et al., 2009; Ritchie & Bates, 2013), which is an essential factor for many health and non-health related decisions (e.g., Dieckmann et al., 2009; Nelson et al., 2008; Reyna et al., 2015; Weinfurt et al., 2003). Enhancing mathematical knowledge, including conceptual and procedural knowledge, is an important educational goal that implies subsequent improvements in various outcomes across the lifespan.

5.2.1 Characteristics of Conceptual and Procedural Knowledge

Conceptual knowledge can be described as the "implicit or explicit understanding of the principles that govern a domain" (Rittle-Johnson et al., 2001, p. 346). It consists of the core concepts (Byrnes & Wasik, 1991) and their interrelations in a knowledge domain. The knowledge is mentally stored as relational representations (e.g., schemas or semantic nets, Hiebert & Lefevre, 1986). Conceptual knowledge can be explicit as well as implicit and, therefore, not always verbalizable (Goldin-Meadow et al., 1993). It develops through the construction of relations between acquired information (Hiebert & Lefevre, 1986), and it often requires knowledge of other concepts from the same or other knowledge domains. This makes conceptual knowledge a multi-dimensional construct (Rittle-Johnson & Schneider, 2014) with rich connections. However, empirical evidence also supports the broader conception of *knowledge of concepts* (Rittle-Johnson & Schneider, 2014). In this definition, the connections between concepts are features of the learners' growing expertise (Schneider & Stern, 2009).

Procedural knowledge is defined as the knowledge of procedures or simply as *knowing how* (Byrnes & Wasik, 1991; Rittle-Johnson et al., 2001). A procedure includes actions that are carried out to accomplish a goal. Subsequently, procedural knowledge differs from other forms of knowledge due to its sequential and goal-directed nature (Hiebert & Lefevre, 1986). Procedural knowledge can be automatized, depending on the extent of the learner's problem-solving practice (Schneider & Stern, 2010). Strongly automatized knowledge is tied to problem types and more difficult to transform (Rittle-Johnson & Schneider, 2014). Further constraints to this definition are bound to the specific knowledge domain. In mathematics, procedural knowledge also includes knowledge about how to carry out procedures (Star, 2005) or to complete tasks (Baroody et al., 2007).

Measures of conceptual and procedural knowledge differ in the domain of mathematics. Researchers implicitly or explicitly assess conceptual knowledge through various tasks, for example, by evaluating examples or providing definitions (Rittle-Johnson & Schneider, 2014). Measures of procedural knowledge are less diverse. In mathematics, problem-solving accuracy on familiar or similar problems is often used (Rittle-Johnson & Schneider, 2014).

One must develop conceptual and procedural knowledge to learn successfully in a domain (e.g., Hurrell, 2021). However, the distinction between these two types of knowledge is complicated by their interlinked nature (Baroody et al., 2007). The dependency between conceptual and procedural knowledge hinders their sole measurement (Rittle-Johnson & Schneider, 2014). On the one hand, problem-solving always requires some procedural knowledge (Rittle-Johnson et al., 2001). On the other hand, when there is only weak procedural knowledge, learners use conceptual knowledge to solve routine problems instead (Braithwaite & Sprague, 2021). Therefore, conceptual and procedural knowledge can hardly be measured entirely independently.

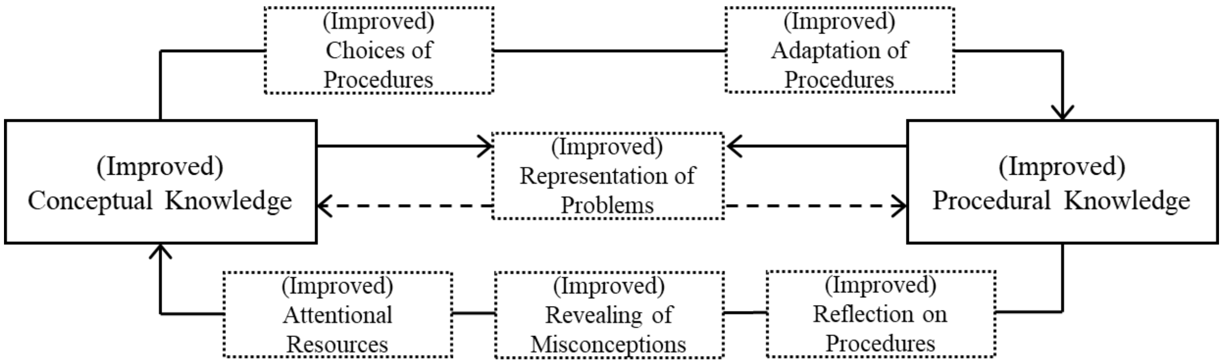
5.2.2 Predictive Relations Between Conceptual and Procedural Knowledge

At least four theoretical viewpoints exist on the longitudinal relations between conceptual and procedural knowledge (e.g., Schneider et al., 2011; Schneider & Stern, 2010). (1) The *Inactivation View* (Haapasalo & Kadjievich, 2000) states that the two knowledge types develop independently from each other (e.g., Nesher, 1986; Resnick & Omansun, 1987). (2) *Concept-First Theories* predicate that conceptual knowledge develops before procedural knowledge in a domain. Subsequently, learners derive procedural knowledge unidirectionally from their concepts through the construction and evaluation of problem-solving processes (e.g., Donlan et al., 2007; Geary, 1994; Gelman, 1993; Gelman & Williams, 1998). (3) *Procedures-First Theories* (Weaver et al., 2018) state that one develops procedural knowledge first and then builds conceptual knowledge on this basis through unidirectional abstraction processes (e.g., Baroody & Gannon, 1984; Baroody & Ginsburg, 1986; Briars & Siegler, 1984; Kerslake, 1986; Siegler & Stern, 1998). (4) The *Iterative Model* (Rittle-Johnson et al., 2001) describes the causal relations between conceptual and procedural knowledge as bidirectional. Therefore, an increase in one of them leads to a subsequent increase in the other (see Figure 6, e.g., Canobi, 2009; Hecht & Vagi, 2012; Rittle-Johnson & Koedinger, 2009; Schneider et al., 2011). According to the Iterative Model, several mechanisms influence the iterative relations between conceptual and procedural knowledge (Rittle-Johnson et al., 2001). First, improved problem representations (Rittle-Johnson et al., 2001), improved choices among alternative procedures (e.g., Crowley et al., 1997), and adaptations of procedures to new problems (e.g., Siegler & Crowley, 1994) are possible explanations of the influence of conceptual on procedural knowledge (Rittle-Johnson et al., 2001). On the other side, also improved problem representations (e.g., Rittle-Johnson & Alibali, 1999), attentional resources (e.g., Geary, 1995), the revealing of misconceptions (e.g., Hiebert, 1992), and reflections on the cause of successful procedures (e.g., Chi et al., 1989) are potential mechanisms of the influence of procedural on conceptual knowledge (Rittle-Johnson et al., 2001).

Previous findings on the predictive relations are inconsistent. All four views on the development of conceptual and procedural knowledge have been supported by empirical evidence. Some studies purported a more substantial influence of conceptual knowledge on procedural knowledge (e.g., Hecht & Vagi, 2010). For example, the National Council of Teachers of Mathematics (NCTM, 2014) explicitly asserted a unidirectional conceptual-to-procedural perspective in their principle. Others found symmetrical relations between both knowledge types (e.g., Schneider et al., 2011).

Figure 6

The Iterative Model for the Development of Conceptual and Procedural Knowledge



Note. Development of conceptual and procedural knowledge according to the Iterative Model with proposed examples of influential mechanisms in dotted boxes (based on Rittle-Johnson et al., 2001, p. 37).

The theoretical views do not contradict each other. However, the Iterative Model is the most accepted and empirically supported view to explain the development of procedural and conceptual knowledge (e.g., Rittle-Johnson et al., 2015). For example, longitudinal studies have shown that equal bidirectional relations accommodate gradual improvements in both knowledge types (Schneider & Stern, 2010) and are still present years later (Cowan et al., 2011). Causal evidence from randomized controlled trials manipulating at least one of the knowledge types also confirmed the Iterative Model (e.g., Canobi, 2009; McNeil et al., 2015; Rittle-Johnson & Alibali, 1999). Further, bidirectional relations generalize across several age groups and subdomains of mathematics (Rittle-Johnson & Schneider, 2014). Despite the large body of studies available to date, it is still open to understanding the symmetrical nature of the bidirectional relations between conceptual and procedural knowledge since results differ across studies.

5.2.3 Moderators of the Relation Between Conceptual and Procedural Knowledge

The influence of moderator variables is one explanation of the four views' coexistence on the relation between procedural and conceptual knowledge and the corresponding empirical variations. Previous literature discussed the learner's age and educational level, the mathematical subdomain, and whether an intervention occurred or not as common moderators.

The strength of the predictive relations between conceptual and procedural knowledge might depend on the age and the attended school level of the learners. There is consensus that conceptual and procedural knowledge are connected across different age groups and educational levels (LeFevre et al., 2006; Jordan et al., 2013; Canobi et al., 2003; Hecht & Vagi, 2012). However, whether these bidirectional predictive relations are symmetrical is still unclear. Previous studies indicated that young learners likely develop procedural knowledge first, facilitating later conceptual understanding (Karmiloff-Smith, 1992; Siegler & Stern, 1998). For example, young children typically learn counting procedures before understanding the underlying concepts (e.g., Frye et al., 1989; LeFevre et al., 2006). On the other side, there exists broad evidence that for older learners, conceptual knowledge is one of the most critical factors for subsequent learning, including the acquisition of procedural knowledge (Hecht et al., 2003; Gelman & Williams, 1998; Halford, 1993; Hiebert & Lefevre, 1986; Schneider et al., 2009). It is argued that conceptual knowledge can be applied to familiar and novel problems. In contrast, procedural knowledge is more tied to a routine (Rittle-Johnson et al., 2001), explaining the prominent role of conceptual knowledge for older (i.e., more knowledgeable) students. The symmetry of the iterative relationship between conceptual and procedural knowledge is likely different due to the learner's age. Once a learner has initial knowledge of a topic, the predictive relations could be more assertive. This meta-analysis investigates whether the predictive relations are symmetrical depending on the learners' age.

Also, this meta-analysis inspects whether the predictive relations get stronger the higher the attended school levels. A reason could be that prior conceptual and procedural knowledge are better developed for children with more expertise. This then functions as a more pronounced starting point for further iterative development. A systematic investigation of the relationship between conceptual and procedural knowledge across age and educational levels is still outstanding.

The bidirectional relationship between conceptual and procedural knowledge may differ in strength due to the subdomains of mathematics. For example, in the subdomain of rational numbers, more precisely for fractions, procedural knowledge is found to be dependent on existing conceptual knowledge (Byrnes & Wasik, 1991). In the subdomain of algebra, prior procedural knowledge is a prerequisite for conceptual knowledge about functions (Lauritzen, 2012). These two examples show that differences in conceptual and procedural knowledge relations between the subdomains of mathematics are possible. While these potential differences have been addressed in previous reviews (Rittle-Johnson & Schneider, 2014; Rittle-Johnson & Siegler, 1998), a synthesis of the quantitative evidence in the form of a meta-analysis is still lacking.

Carrying out an intervention might moderate the predictive relations between conceptual and procedural knowledge. The development of both knowledge types can be experimentally enhanced through instructional interventions promoting one or both (e.g., Blöte et al., 2001; Canobi, 2009; Hiebert & Wearne, 1996; McNeil et al., 2012). Based on the Iterative Model, interventions that address only one of the knowledge types lead to subsequent improvements in the other as well (Rittle-Johnson & Alibali, 1999). Therefore, the predictive relations between conceptual and procedural knowledge might be more assertive when an intervention takes place; regardless of whether it aims to enhance one or both knowledge types.

Other exploratory moderators have not yet been examined in depth in previous research. Therefore, this meta-analysis investigates the possible influences of time lags between measurements as well as the broad and specific task types of the assessment of conceptual knowledge.

Time lags between measurements influence the investigated effects. They are an essential aspect of its understanding (Gollob & Reichardt, 1987) and should be determined by the change process of the observed constructs (Little, 2013). According to this, the strength of the predictive relations should decrease when the time lags increase. However, the bidirectionality of the predictive relations between conceptual and procedural knowledge should not differ in dependence on time lags.

Additionally, the task types of the assessment of conceptual knowledge might influence the predictive relations between conceptual and procedural knowledge. Rittle-Johnson and Schneider (2015) differentiated between two broad task types and 13 specific task types for the assessment of conceptual knowledge. According to them, conceptual knowledge can be measured implicitly, for example, by qualitatively evaluating or rating answers given by others. Explicit measures, on the other side, include explanations or generating definitions, among other tasks. The broad as well as specific task types might influence the predictive relations between conceptual and procedural knowledge. The more reliable a task captures conceptual knowledge, the better conceptual and procedural knowledge can be differentiated. Because procedural knowledge usually is measured by the accuracy of responses, we did not differentiate task types for it.

5.2.4 Investigating Predictive Relationships: The Use of Cross-Lagged Panel Models

In the research on conceptual and procedural knowledge, cross-lagged panel models (CLPM) are widely used because of the possibility of investigating the cross-lagged effects while controlling for the stability of the included constructs (Hamaker et al., 2015). A CLPM provides information on rank-order stability and how prior scores of one variable relate to changes in the other variable (Mund & Nestler, 2019). The latter can be interpreted as predictive relations over time between both variables (Berry & Willoughby, 2017).

Many primary studies use the approach of CLPM because of practical reasons, although the CLPM makes some assumptions that can be criticized in applied contexts (Hamaker et al., 2015; Lucas, 2023; Mund & Nestler, 2019): For example, the conjecture that all participants vary across one group mean without between-person differences in the variables is problematic (Mund & Nestler. 2019). Second, between- and within-person relations are not distinguishable in the parameters of a CLPM (Berry & Willoughby, 2017; Mund & Nestler, 2019) and may therefore be conflated (Lucas, 2023). Third, the CLPM does not capture mean-level changes in the observed constructs over time and whether persons differ in this change (Mund & Nestler, 2019).

However, CLPM is still the standard for examining predictive relations between two variables. In contrast to other more advanced models, two waves of data are enough for a CLPM (Kenny, 2005). This makes it a useful approach from a practical perspective (Hamaker et al., 2015). Additionally, from the statistical point of view, CLPMs are more suited to evaluate possible causal relations than simple cross-lagged correlations or multiple regressions (Hertzog & Nesselroade, 2003). This advantage is premised on the autoregressive effects of the two variables in a CLPM which enable controlling for their influence on the cross-lagged paths (Orth et al., 2021).

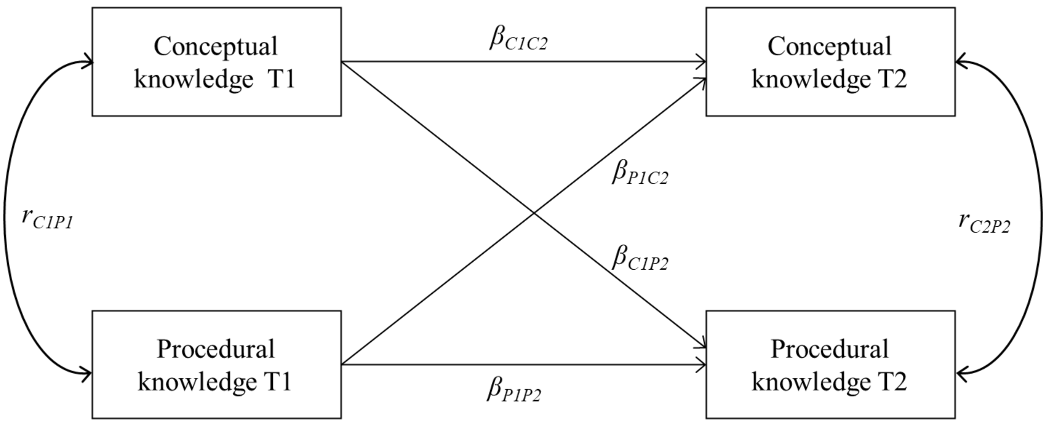
We examined the predictive relations between conceptual and procedural knowledge with a CLPM for two reasons: First, most of the primary studies up to this point provided data that could be included in a CLPM. Second, when investigating the predictive relations between conceptual and procedural knowledge, all four theoretical views on the pattern of relationships can be explored using a CLPM (see Figure 7). According to the Inactivation View, we would assume that the cross-lagged paths β_{C1P2} and β_{P1C2} are nonsignificant. A more substantial effect of β_{C1P2} and a nonsignificant effect of β_{P1C2} would support the Concepts-First Theories. Conversely, a significant β_{P1C2} and a nonsignificant β_{C1P2} would underpin the Procedures-First Theories. Last, the Iterative Model expects both cross-lagged paths (β_{C1P2} and β_{P1C2}) to be significant and symmetrical in the CLPM.

5.3 The Current Study

The current meta-analysis aims to shed light on the predictive relations between conceptual and procedural knowledge considering the four theoretical views and several moderator variables discussed in the literature. While previous studies provided empirical

Figure 7

Cross-Lagged Panel Model of Conceptual and Procedural Knowledge at Two Measurement Points



Note. Cross-lagged panel model of conceptual knowledge at measurement point one (C1), procedural knowledge at measurement point one (P1), conceptual knowledge at measurement point two (C2), procedural knowledge at measurement point two (P2) as well as the cross-sectional (r_{C1P1} and r_{C2P}), auto-regressive (β_{C1C2} and β_{P1P2}) and cross-lagged paths (β_{C1P2} and β_{P1C2}) between the four variables.

evidence for each of the four viewpoints on the relationship between conceptual and procedural knowledge, it is still open how this relation is moderated by third variables such as learner characteristics, characteristics of the knowledge domain, or whether an intervention took place.

To comprehensively understand the longitudinal relationship between conceptual and procedural knowledge, we specified a CLPM as a meta-analytic structural equation model (MASEM) to examine the predictive relations between both knowledge types. Using a MASEM, the structural equation model with all its path coefficients can be inspected simultaneously in the meta-analysis. We propose the following five hypotheses:

(1) The predictive relations between conceptual and procedural knowledge are bidirectional and symmetrical. Following the Iterative Model, we expect significant medium positive predictive relations between conceptual knowledge at measurement point one (T1) and procedural knowledge at measurement point two (T2) and vice versa. This means the

path coefficients β_{C1P2} and β_{P1C2} are significant, medium strong, and symmetrical in the CLPM.

(2) The learners' age moderates the predictive relations between conceptual and procedural knowledge. We expect that age has an impact on the symmetry of the iterative relationship so that procedural knowledge at T1 is more strongly associated with conceptual knowledge at T2 ($\beta_{C1P2} < \beta_{P1C2}$) for young learners and that conceptual knowledge at T1 is a better predictor of procedural knowledge at T2 for older, more knowledgeable learners ($\beta_{C1P2} > \beta_{P1C2}$). Including age as a moderator adds information on individual differences that are not fully represented when only including the individual school level since formal schooling can start at different ages across countries. Further, as a continuous moderator, considering the learners' mean age makes examinations of linear trends feasible.

(3) The predictive relations between conceptual and procedural knowledge depend on the individual's school level. We expect that learners at lower school levels have not yet fully developed both types of knowledge and differ in their knowledge profiles (Hallett et al., 2010; Hecht & Vagi, 2012). Additionally, higher school-level students might already possess more prior conceptual and procedural knowledge. They are better suited to further improve both knowledge types on that basis. We assume the higher the school level, the stronger the predictive relations between conceptual and procedural knowledge (β_{C1P2} and β_{P1C2}). With the school level as a categorical moderator, it is possible to examine nonlinear trends.

(4) The predictive relations between conceptual and procedural knowledge differ between subdomains of mathematics. There are different findings on whether conceptual knowledge predicts procedural knowledge (Byrnes & Wasik, 1991) or vice versa (Lauritzen, 2012) in dependence on the subdomain of mathematics. Therefore, we

exploratorily analyze differences in the relations (β_{C1P2} and β_{P1C2}) due to the broad subdomains of mathematics.

(5) The predictive relations between conceptual and procedural knowledge are stronger in studies where an intervention occurred than in those without an intervention. Interventions can strengthen conceptual as well as procedural knowledge (e.g., Canobi, 2009; McNeil et al., 2012) and their relations (Blöte et al., 2001; Hiebert & Wearne, 1996). As such, we expect stronger predictive relations between conceptual and procedural knowledge (β_{C1P2} and β_{P1C2}) for intervention studies than those without any instruction.

In addition, we conducted several exploratory moderator analyses. These moderators were the time lag between measurements and the broad and specific task types of conceptual knowledge. An overview of the included moderators is provided in Table 1.

Table 1

Description of the Included Moderators (Study 1)

Moderator	Description
Knowledge characteristics	
Subdomain of mathematics	To test whether the relation between conceptual and procedural knowledge at T1 and T2 differs between studies analyzing knowledge in different subdomains of mathematics, we coded five subdomains: *Counting and Whole-Number Arithmetic*, *Rational Numbers*, *Algebra*, *Other*, and *Several*.
Sample characteristics	
Mean age	We coded the sample mean age of the learners (in years) at T1.
School level	We distinguished between different school levels. We coded the following levels: *Kindergarten and Preschool*, *Primary School*, *Secondary School*, *Higher Education*, *Continued Education*, and *Several*.
Study characteristics	
Time lag between measurements	The time lag between measurements might moderate the relation between conceptual and procedural knowledge at T1 and T2. We coded the time lag in days.

Table 1 (continued)

Moderator	Description
Intervention	To test whether the relation between conceptual and procedural knowledge at T1 and T2 differs between studies with and without interventions, we coded the two levels, *Intervention* and *No Intervention*.
Methodological characteristics	
Broad task of conceptual knowledge	We differentiated between *Explicit* and I*mplicit* broad task types for the assessment of conceptual knowledge as well as *Both* task types.
Specific task of conceptual knowledge	As specific task types of the assessment of conceptual knowledge, we coded 13 levels: *Evaluate Unfamiliar Procedures*, *Evaluate Examples of Concept*, *Evaluate Quality of Answers Given by Others*, *Translate Quantities Between Representational Systems*, *Compare Quantities*, *Invent Principle-Based Shortcut Procedures*, *Encode Key Features*, *Sort Examples Into Categories*, *Explain Judgments*, *Generate or Select Definitions of Concepts*, *Explain Why Procedures Work*, *Draw Concept Maps*, and *Several*.

Note. Description of included moderators and moderator levels.

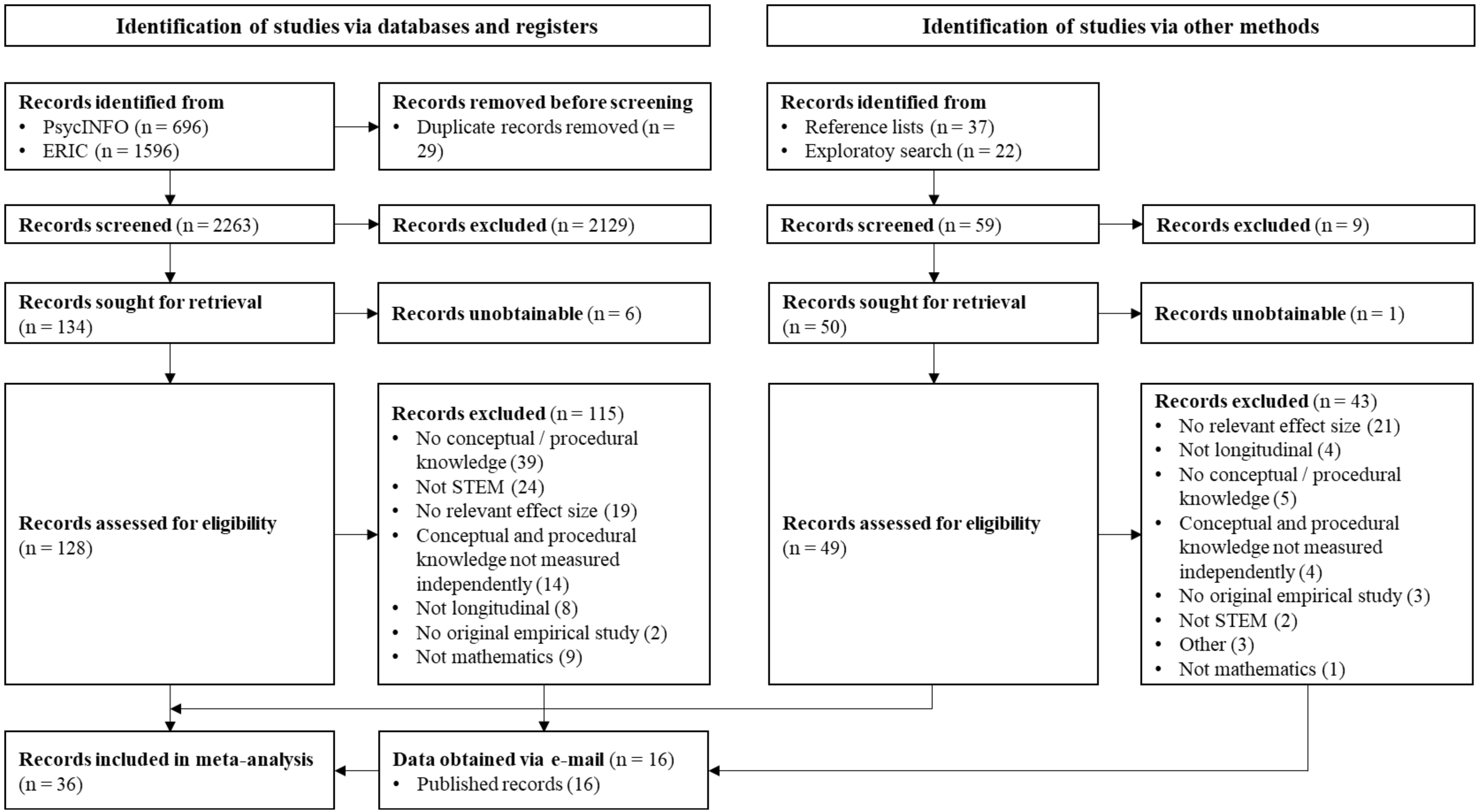
5.4 Method

5.4.1 Literature Search

We conducted a standardized literature search in PsycINFO and ERIC on 21 October 2020, as well as an exploratory literature search in Google Scholar and reference lists (Rittle-Johnson & Schneider, 2014; Rittle-Johnson et al., 2015; Simonsmeier et al., 2022). The search string in PsycINFO was (concept* or principle*) and (procedur* or skill*) in abstract and title with the limits set to human, non-disordered population, English language, and longitudinal study. Because of fewer available limits in ERIC, the search string contained more specifications: (((concept* or principle*) and (procedur* or skill*) and (develop* or longitudinal* or predict* or iterativ*) and (empiric* or measure* or data* or tested*) and knowledge*) NOT review*) in abstract and title. Here, the limits were set to journal articles or dissertations/theses - doctoral dissertations. The standardized literature search provided 2292 studies in total. The exploratory search returned 59 additional studies. Figure 8 summarizes the literature search process in a PRISMA flow diagram (Page, McKenzie, et al., 2021).

Figure 8

Flow Chart for the Literature Search and Inclusion Process (Study 1)



Note. Flow diagram of the study search and inclusion process (adapted from Page, McKenzie, et al., 2021; Page, Moher, et al., 2021). The process on the left is based on records identified by the systematic search; the process on the right is based on the exploratory search.

5.4.2 Inclusion of Studies

The first inclusion criterion for studies was (a) the study included two or more measurement points with assessments of learners' conceptual and procedural knowledge in STEM domains. We included studies with correlation matrices that provided at least one measurement point with measurements of both knowledge types and a second measurement point with at least one of the types. The measurements had to be objective quantitative measures of domain-specific knowledge. Combined scores of conceptual and procedural knowledge as well as measures of meta-cognitive knowledge, intelligence, achievement, basic cognitive functions, self-assessments, and other domains than STEM domains were excluded. We only included studies in which conceptual and procedural knowledge were measured in the same domain. Further inclusion criteria were (b) the study included assessments of the longitudinal bivariate correlations between conceptual and procedural knowledge and (c) the study provided the information necessary to compute standardized effect sizes, which could be converted into correlations.

We excluded studies that used measurements with time limits (e.g., Bailey et al., 2012) because time limits might restrict the potential knowledge shown in assessments. We also excluded studies with solution time as a measure (e.g., Paul et al., 2019) and those with measurements that were not clearly assignable to either conceptual or procedural knowledge (e.g., Matthews & Fuchs, 2020).

We determined study eligibility in two steps: First, the first author screened titles and abstracts of the search hits, and the studies were either excluded with reasons or retained for inspection of the full texts. A research assistant double-coded 100 abstracts independently after receiving coder training with the first author. The agreement was 82%, and disagreements were resolved by discussion. In the second step, the full texts were screened and excluded if they failed to meet the inclusion criteria. After a coder training, the first author and a co-author independently coded 100 studies for eligibility. The mean interrater agreement for this step was

98%. Overall, 22 studies could not be included because the relevant information was neither provided in the study nor by the authors upon email request. We received one additional manuscript under review at the time of inclusion (Gunderson & Hildebrand, 2021).

In the study selection process, a restriction occurred in the availability of studies within the STEM domains other than mathematics. Only two studies with enough information available to be included in the meta-analysis were outside of the domain of mathematics. Hence, we decided to limit the analysis to the knowledge domain of mathematics. A final number of 552 predominantly dependent correlation matrices, called effect sizes in the following, from 36 studies were aggregated to 60 independent effect sizes on sample level and included in this meta-analysis at this point.

5.4.3 Data Coding

Data coding was based on a coding manual. This manual was used for coder training and coding of the articles. The following information was gathered for all included studies: We coded study characteristics, such as the authors, the publication year, if the article was peer-reviewed, the literature type, the sample size, the study design, the total number of measurement points, and which of these measurement points were used for this meta-analysis. We only coded each study's first and last possible measurement point to minimize the number of dependent effect sizes in the meta-analysis. Additionally, we coded the setting of the study, the time lag between used measurement points, if there was an intervention between measurements, the group size of the instruction, and the group size of the knowledge assessments. Coded information about the knowledge domain was the subdomain in mathematics. Sample characteristics were the school level at T1, the school level at T2, the class at T1, the class at T2, the gender of the participants, the age group at T1, the age group at T2 as well as the mean age at T1 and the mean age at T2. Also, the country and the continent of the primary study were coded. The conceptual and procedural knowledge measures at T1 and T2 included whether the measurement was self-constructed, the broad and specific task type for conceptual knowledge, the name of the used test or scale, the measure type, and the reliability. As effect sizes, we coded all six bivariate correlations between the two types of knowledge measured at two measurement points: Conceptual and procedural knowledge at T1 (C1P1), conceptual knowledge at T1 and T2 (C1C2), conceptual knowledge at T1 and procedural knowledge at T2 (C1P2), procedural knowledge at T1 and conceptual knowledge at T2 (P1C2), procedural knowledge at T1 and T2 (P1P2), and conceptual and procedural knowledge at T2 (C2P2). Of the six possible correlations, all studies reported at least three. Missing and unclear information was coded as missing.

In the third step, the first author and a co-author coded the study characteristics for the 64 effect sizes that were available at that time. The mean interrater agreement was 95%. The agreement ranged from 83% to 100% for the coded variables, except for the specific type of conceptual knowledge at T1. For this variable, the agreement was 40%, mainly because the two raters disagreed on one study with several effect sizes (without this study, the agreement would have reached 76%). Therefore, we retained this variable. Additionally, the corresponding authors of 41 studies with missing or unclear information were contacted to request additional data. Ten authors provided the requested data (response rate: 33,33%). The first author coded the studies with additional information. Also, we conducted a Risk of Bias assessment for all included primary studies. This Risk of Bias assessment was based on RoB2 (Sterne et al., 2019) and modified for our purposes (see Tables 18 and 19 in Appendix E).

5.4.4 Preparation of Effect Sizes

We used Pearson correlation (*r*) as the effect size of the meta-analysis. All included studies reported correlations, so none of the effect sizes had to be converted. For each study, we put the up to six correlations in a correlation matrix. We visualized the coded effect sizes using forest plots. By interpreting these forest plots, one must consider the differences between the altitude of the single effect sizes and the path coefficients of the CLPM. This difference results from the feature of CLPM to control for the stability of the included constructs (Hamaker et al., 2015).

We tested for outliers by inspecting Cook's values (Cook & Weisberg, 1982) using the *metafor* package (Viechtbauer & Viechtbauer, 2015). Cook's values, also called Cook's distance, measure the influence of each effect size in the model on the main result. This is done by removing each effect size from the model and summarizing the change this makes in the resulting model (Viechtbauer & Cheung, 2010). Through this process, outliers that excessively influence the model and need further inspection can be identified. Rules of thumb promote several possible cut-off points. We removed effect sizes with a Cook's distance greater than four divided by the total number of effect sizes. This applied to nine independent correlation matrices from six studies (Cowan et al., 2011; Ngu & Phan, 2016; Rittle-Johnson & Koedinger, 2009; Rittle-Johnson et al., 2016; Schneider et al., 2011; Wasiu & Abiola, 2019).

5.4.5 Statistical Analysis

5.4.5.1 Publication Bias. We visually and statistically tested for publication bias using funnel plots and Egger regressions (Egger et al., 1997). These tests were conducted using the *metafor* package (Viechtbauer & Viechtbauer, 2015) in R (R Core Team, 2022) for each of the six bivariate correlations.

5.4.5.2 Meta-Analytic Integration. We used a one-stage meta-analytic structural equation model (OSMASEM) for the analysis. This model allowed us to meta-analyze all cross-lagged panel paths in a single model. Additionally, it enabled us to test moderator hypotheses with categorical and continuous moderators (Cheung, 2014). In a MASEM, structural equation models (SEM) are fitted to meta-analytic data (Jak et al., 2021). With this, a MASEM combines the strengths of both meta-analyses and SEM (Cheung, 2015b). Typically, a MASEM consists of two stages (Cheung & Chan, 2005): In the first stage, the correlation matrices from primary studies are all combined into one pooled correlation matrix. This matrix is then fitted by a SEM in the second stage (Jak & Cheung, 2020). In contrast, in an OSMASEM, the SEM is directly fitted to the set of correlation matrixes of the primary studies. We conducted the analyses using the *metaSEM* package (Cheung, 2015a) and the *semPlot* package (Epskamp, 2019) in R (R Core Team, 2022). Heterogeneity, in the form of between-studies variance, was determined by extracting the heterogeneity variance-covariance matrix with the *metaSEM* package (Cheung, 2015a). The data and statistical code used in this meta-analysis are available upon request.

Most of the included studies reported several effect sizes for the same sample; for example, for various measures of conceptual and procedural knowledge. These effect sizes are statistically dependent, violating classic meta-analytical models' central assumption (e.g., Cheung, 2019). So far, it is impossible to account for dependent effect sizes in OSMASEM. Therefore, we averaged the effect sizes on sample level. Five primary studies used the same sample for their analyses (Bailey et al., 2017; Hansen et al., 2015; Hansen et al., 2017; Jordan et al., 2013; Ye et al., 2016). We aggregated the effect sizes from these studies to one mean effect size for each of the six bivariate correlations.

5.4.5.3 Moderators. One advantage of OSMASEM is the possibility of examining both categorical and continuous moderators (Jak & Cheung, 2020). These moderator analyses are based on Bauer's (2017) method to model SEM parameters as a function of a moderator variable (Jak & Cheung, 2020). The effect sizes were averaged on moderator level within the samples. The moderators *School Level*, *Subdomains Mathematics*, *Intervention*, *Broad Task Type*, and *Specific Task Type* were specified as categorical moderators. A comparison of more than two levels of a categorical moderator is impossible in OSMASEM. Therefore, we used pairwise comparisons. The continuous moderators *Mean Age* at measurement point one and the *Time Lag* between the measurements were z-standardized to facilitate the interpretation of the results (Jak et al., 2021). An overview of subgroup results for the moderator levels can be found in Table 17 in Appendix B.

5.5 Results

5.5.1 Publication Bias

Figures 35 to 40 in Appendix C display the funnel plots for each of the six correlations in the cross-lagged panel model. The plots show no indication of a publication bias. In line with this, the Egger regressions did not reach significance for any of the six correlations. This means there is no underrepresentation of effect sizes close to zero in the included data.

5.5.2 Study Characteristics

The analyzed data comprised 51 independent correlation matrices from 34 studies with a total of 7253 participants. The median sample size for the correlation matrices was 62, and the sample mean age ranged from 3.85 to 16.33. At the first measurement point, the participants were in kindergarten or preschool in seven studies, in primary school in 14 studies, in secondary school in nine studies, and in several school levels in two studies. Twenty-three studies were conducted in North America, followed by eight from Europe and three from Asia.

The investigated subdomain was counting and whole-number arithmetic in 11, algebra in 10, and rational numbers in four studies. One study was conducted in another subdomain (proportional reasoning, Noche et al., 2019), and nine used contents from several subdomains. Eighteen studies did not have an intervention, and interventions were conducted in the other 14 studies. The results of a Risk of Bias assessment of all included studies can be found in Appendix E (Tables 20 and 21). Table 2 presents an overview of the included studies (for the list of included studies, see Appendix A).

Table 2

Overview of Included Studies (Study 1)

Study	Sample	*j*	*N*	Mean age	School level	Time lag	Subdomain	C1P1	C1C2	C1P2	P1C2	P1P2	C2P2
Achmetli et al. (2019)	1	1	97	16.33	Secondary school	0	Algebra	.304	.260	.389	.384	.531	.510
	2	1	101	16.33	Secondary school	0	Algebra	.411	.515	.402	.535	.513	.448
	3	1	89	16.33	Secondary school	0	Algebra	.476	.411	.372	.613	.450	.567
Bailey et al. (2014)	1	1	37	6.92	Primary school	2920	Several	.550	.620	-	.560	-	-
Bailey et al. (2017)[a]		1	481	-	Primary school	639	Several						
Hansen et al. (2015)[a]		5	334	10.82	Primary school	90 - 456	Several						
Hansen et al. (2017)[a]	1	1	481	-	Primary school	1186	Algebra	.406	.507	.354	.458	.481	.547
Jordan et al. (2013)[a]		2	357	8.83	Primary school	91 - 456	Several						
Ye et al. (2016)[a]		3	481	10.83	Primary school	274 - 365	Several						
Byrnes (1992)	1	4	25	12.58	Secondary school	1	Whole-number	.550	-	.505	-	.400	-
Ching (2016)	1	30	115	6.36	Primary school	300	Whole-number	.395	-	.477	-	.523	-
Chu et al. (2016)	1	4	100	3.91	Kindergarten / Preschool	365	Whole-number	.240	.375	.410	.270	.360	.460
Cowan et al. (2011)	1[b]	4	259	7.92	Primary school	365	Whole-Number	.735	.675	.705	.685	.880	.700
Fyfe et al. (2014)	1	1	60	8.20	Primary school	14	Algebra	.331	.648	.323	.300	.179	.714
	2	1	62	8.20	Primary school	14	Algebra	.101	.573	.321	.237	.299	.565
Geary et al. (2019)	1	3	197	3.83	Kindergarten / Preschool	488	Whole-Number	.273	.206	-	.183	-	-
Gunderson & Hildebrand (2021)	1	1	283	6.74	Several	548	Whole-Number	.439	.535	.382	.429	.477	.435
Hecht & Vagi (2010)	1	144	181	-	Primary school	456	Several / Rational numbers	.464	.593	.519	.456	.498	.527

Table 2 (continued)

Study	Sample	*j*	*N*	Mean age	School level	Time lag	Subdomain	C1P1	C1C2	C1P2	P1C2	P1P2	C2P2
Heemsoth & Heinze (2016)	1	1	87	13.02	Secondary school	60	Rational numbers	.365	.377	.273	.649	.658	.762
	2	1	87	13.02	Secondary school	60	Rational numbers	.594	.672	.510	.634	.790	.601
Jõgi & Kikas (2016)	1	4	864	7.46	Primary school	730	Several	.220	-	.305	-	.345	-
Koponen et al. (2007)	1	1	178	6.25	Kindergarten / Preschool	1670	Whole-Number	.700	-	.480	-	.500	-
Koponen et al. (2016)	1	2	378	6.17	Kindergarten / Preschool	365	Whole-Number	.300	-	.230	-	.515	-
Laski et al (2016)	1	1	15	6.17	Kindergarten / Preschool	810	Whole-Number	.458	.725	.600	.394	.426	.834
	2	1	8	6.17	Kindergarten / Preschool	810	Whole-Number	.924	.552	.948	.522	.799	.713
	3	1	17	7.17	Primary school	780	Whole-Number	.204	.030	.220	.302	.409	.718
	4	1	13	7.17	Primary school	780	Whole-Number	.458	.560	.224	.505	.317	.491
LeFevre et al. (2013)	1	1	157	9.16	Primary school	365	Whole-Number	.520	.590	.470	.420	.520	.490
Ngu & Phan (2016)	1	1	32	12.91	Secondary school	0	Algebra	-	-	-	.060	.667	.234
	2 [b]	1	30	12.91	Secondary school	0	Algebra	-	-	-	.226	.876	.388
Noche et al. (2019)	1	1	19	16.33	Several	17	Other	.509	.770	.611	.702	.835	.812
	2	1	19	16.33	Several	17	Other	.461	.714	.305	.356	.654	.297
	3	1	40	16.33	Several	17	Other	.344	.762	.286	.440	.652	.388
Östergren & Träff (2013)	1	8	305	6.62	Kindergarten / Preschool	365	Whole-number	.593	-	.440	-	.469	-
Rittle-Johnson & Alibali (1999)	1	4	17	10.17	Primary school	1	Algebra	-	.195	.141	-	-	.374
	2	4	16	10.17	Primary school	1	Algebra	-	.412	.258	-	-	.441

Table 2 (continued)

Study	Sample	*j*	*N*	Mean age	School level	Time lag	Subdomain	C1P1	C1C2	C1P2	P1C2	P1P2	C2P2
	3	4	15	10.17	Primary school	1	Algebra	-	.182	.267	-	-	.475
Rittle-Johnson & Koedinger (2009)	1	2	43	11.00	Secondary school	0	Rational numbers	.234	.325	.116	.199	.246	.233
	2 [b]	2	44	11.00	Secondary school	0	Rational numbers	.127	.376	.222	.228	.359	.126
	3	2	13	-	Secondary school	0	Rational numbers	.383	.385	.189	.438	.295	.451
	4	2	13	-	Secondary school	0	Rational numbers	.542	.592	.167	.359	.103	.367
Rittle-Johnson & Star (2009)	1	1	54	13.10	Secondary school	18	Algebra	.718	.612	.459	.522	.457	.751
	2	1	56	13.10	Secondary school	18	Algebra	.401	.628	.474	.506	.519	.579
	3	1	52	13.10	Secondary school	18	Algebra	.378	.691	.542	.429	.324	.715
Rittle-Johnson et al. (2011)	1	1	62	14.10	Secondary school	29	Algebra	.048	.481	.339	.305	.295	.657
	2	1	67	14.10	Secondary school	29	Algebra	.406	.449	.368	.390	.553	.589
	3	1	69	14.10	Secondary school	29	Algebra	.259	.454	.155	.494	.404	.538
Rittle-Johnson et al. (2016)	1 [b]	1	46	7.60	Primary school	26	Algebra	.773	.739	.637	.764	.735	.747
	2 [b]	1	45	7.60	Primary school	26	Algebra	.527	.843	.634	.645	.689	.790
	3 [b]	1	43	7.60	Primary school	26	Algebra	.558	.835	.502	.628	.847	.637
	4 [b]	1	46	7.60	Primary school	26	Algebra	.620	.830	.575	.583	.586	.663
Rittle-Johnson et al. (2017)	1	24	519	4.44-7.00	Kindergarten / Preschool / Primary school	1460-1825	Several	.588	.412	-	504	-	-
Schneider & Stern (2010)	1	64	31	11.20	Primary school	6	Rational numbers	.445	.440	.338	.458	.445	.318
	2	64	27	11.40	Primary school	6	Rational numbers	.205	.415	.308	.291	.366	.437

Table 2 (continued)

Study	Sample	*j*	*N*	Mean age	School level	Time lag	Subdomain	C1P1	C1C2	C1P2	P1C2	P1P2	C2P2
	3	64	26	11.30	Primary school	6	Rational numbers	.382	.420	.243	.284	.245	.507
	4	64	205	11.30	Primary school	120	Rational numbers	.303	.368	.260	.305	.278	.449
Schneider et al. (2011)	1	1	228	13.3	Secondary school	5	Algebra	.191	.440	.381	.336	.366	.715
	2 [b]	1	304	14.00	Secondary school	5	Algebra	.687	.702	.680	.643	.725	.733
Star et al. (2015)	1	1	762	-	Secondary school	180	Algebra	.407	.489	.420	.335	.379	.719
	2	1	570	-	Secondary school	180	Algebra	.395	.518	.469	.331	.364	.727
Stock et al. (2007)	1	1	108	5.90	Kindergarten / Preschool	365	Whole-Number	.600	.190	.320	.160	.370	.160
Vukovic et al. (2014)	1	1	163	7.92	Primary school	730	Several	.460	.520	-	.580	-	-
Wasiu & Abiola (2019)	1 [b]	1	92	-	Higher education	42	Algebra	.720	.090	.010	.131	.037	.250
	2 [b]	1	90	-	Higher education	42	Algebra	.857	-.167	-.143	-.123	-.151	.741

Note. Sample - Number of independent samples in the primary study, *j* - Number of extracted correlation matrices, *N* - Total *N* of the study, Mean age - Mean age of participants at measurement point one, School level - School level at measurement point one, Time lag - Time between measurements in days, Subdomain - Subdomain in mathematics, C1P1 - Aggregated correlation conceptual knowledge and procedural knowledge T1, C1C2 - Aggregated correlation conceptual knowledge T1 and T2, C1P2 - Aggregated correlation conceptual knowledge T1 and procedural knowledge T2, P1C2 - Aggregated correlation procedural knowledge T1 one and conceptual knowledge T2, P1P2 - Aggregated correlation procedural knowledge T1 and T2, C2P2 - Aggregated correlation conceptual knowledge and procedural knowledge T2.

[a] These primary studies have been aggregated to mean correlations because they were conducted on the same sample.

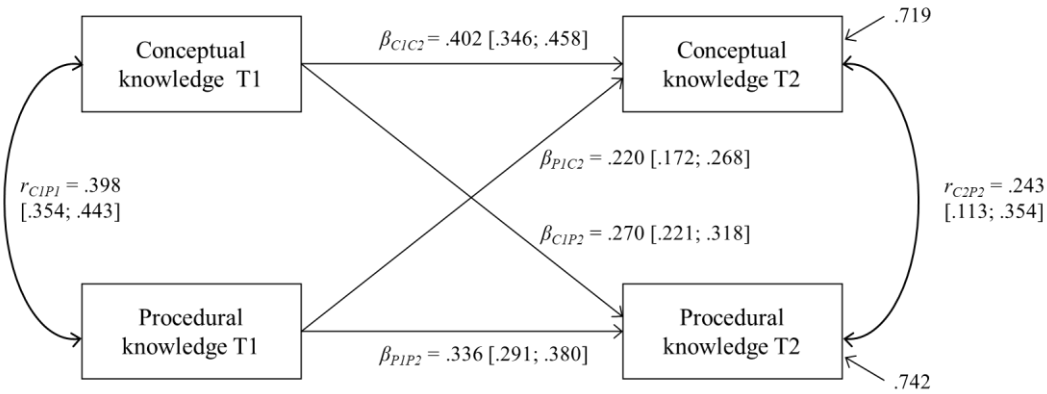
[b] These effect sizes were marked as outliers based on Cook's distance and were not included in the statistical analyses.

5.5.3 Main Meta-Analytic Results

5.5.3.1 Predictive Relations Between Conceptual and Procedural Knowledge. As expected, all standardized path coefficients in the CLPM were positive (Figure 9). The cross-lagged paths (β_{C1P2} and β_{P1C2}) carry information on the predictive relations between conceptual and procedural knowledge. These standardized path coefficients show the influence of one knowledge type at T1 on the change of the other from T1 to T2. Simultaneously, the autoregressive effects of each knowledge type on itself are controlled for.

Figure 9

Cross-Lagged Panel Model of Overall Results



Note. Cross-lagged panel model of conceptual knowledge at measurement point one (C1), procedural knowledge at measurement point one (P1), conceptual knowledge at measurement point two (C2), procedural knowledge at measurement point two (P2) as well as the cross-sectional (r_{C1P1} and r_{C2P}), auto-regressive (β_{C1C2} and β_{P1P2}) and cross-lagged paths between the four variables (β_{C1P2} and β_{P1C2}); k = 41 to 47 for the six bivariate correlations.

The predictive relation of conceptual knowledge at T1 to procedural knowledge at T2 was statistically significant (β_{C1P2} = .270, 95% CI [.221, .318]). This relation describes the prediction of individual differences in conceptual knowledge at T1 on the change in individual differences in procedural knowledge from T1 to T2. The predictive relation of procedural knowledge at T1 with conceptual knowledge at T2 was slightly less strong but statistically significant (β_{P1C2} = .220, 95% CI [.172, .268]). This path coefficient indicates how well individual differences in procedural knowledge at T1 predict the change in individual differences in conceptual knowledge from T1 to T2. According to Orth and colleagues (2022), both cross-lagged effects are large. The forest plots for each of the six bivariate correlations can be found in Appendix D (Figures 41 to 46).

This finding confirms the main hypothesis H1 that there are similar significant positive predictive relations in the cross-lagged paths between conceptual knowledge at T1 and procedural knowledge at T2 and vice versa.

5.5.3.2 Heterogeneity. The between-studies variance (τ^2) is identified by extracting the variance-covariance matrix of the random effects of the overall model. It ranged from .005 to .129. The between-studies variance indicated that heterogeneity existed in this meta-analysis and that there might be influences of moderator variables.

Table 3

Between-study Variance of Cross-lagged Panel Model

τ^2_1	τ^2_2	τ^2_3	τ^2_4	τ^2_5	τ^2_6
.015	.013	.007	.007	.005	.129

Note. τ^2_1 to τ^2_6 - Heterogeneity index in the form of between-study variance for each of the six path coefficients.

5.5.4 Moderator Analyses

5.5.4.1 Age. To test hypothesis H2, the mean age influences the symmetry of the iterative relationship; we conducted a moderator analysis of the continuous moderator *Mean Age* at measurement point one. We expected procedural knowledge at T1 to be more strongly associated with conceptual knowledge at T2 ($\beta_{C1P2} < \beta_{P1C2}$) for young learners and conceptual knowledge at T1 to be a better predictor of procedural knowledge at T2 for older learners (β_{C1P2}

> β_{P1C2}). The results of this analysis must be interpreted with caution because the Hessian was not convex, and the information matrix was not positive definite for this model. The predictive relation from conceptual to procedural knowledge changed by β_{C1P2} = .010 (95% CI [-; -]) and from procedural to conceptual knowledge by β_{P1C2} = -.003 (95% CI [-; -]) when the mean age increased. Both path coefficients did not show significant changes. We did not find a moderating effect of the mean age, and hypothesis H2 was not confirmed.

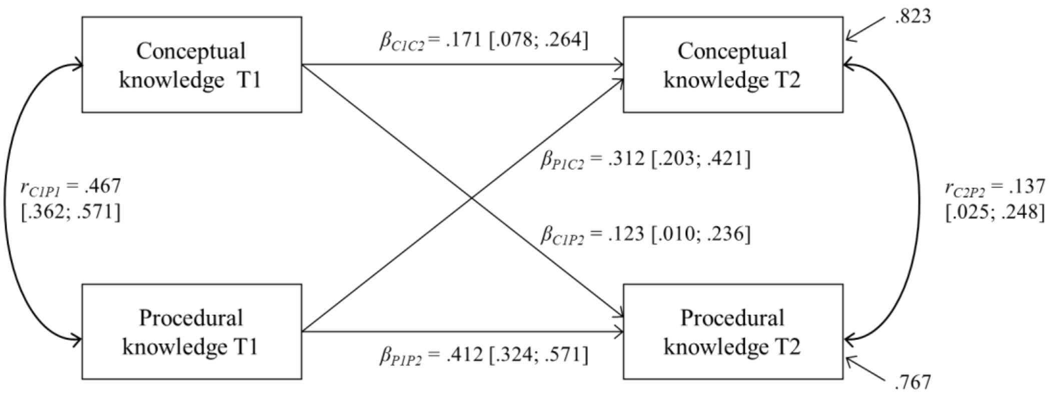
5.5.4.2 School Level. The categorical moderator *School Level* initially contained six levels. In the following moderator analysis, only three levels, *Kindergarten / Preschool*, *Primary School*, and *Secondary School*, are compared. We did not examine the levels *Higher Education*, *Continued Education*, and *Several* because they each only included two or fewer independent correlation matrices. According to hypothesis H3, we tested whether the predictive relations between conceptual and procedural knowledge (β_{C1P2} and β_{P1C2}) were stronger for higher school levels than for lower.

The CLPMs for the three moderator levels are presented in Figures 10 to 12. The path coefficients β_{C1P2} and β_{P1C2} were statistically significant for all moderator levels. The path coefficient β_{C1P2} was descriptively weaker than the path coefficient β_{P1C2} for the level *Kindergarten / Preschool*. Contrary, for the other two moderator levels, the predictive relation from conceptual to procedural knowledge (β_{C1P2}) was descriptively stronger than vice versa. The results of the model *Kindergarten / Preschool* must be interpreted with caution because the Hessian at the solution was not convex and the information matrix was not positive definite.

The levels *Kindergarten / Preschool*, *Primary School*, and *Secondary School* were compared pairwise in separate moderator analyses. No significant differences could be found between these levels in the predictive relations between conceptual and procedural knowledge. The changes in the path coefficients for comparing the models are presented in Table 4. Hypothesis H3 was not confirmed.

Figure 10

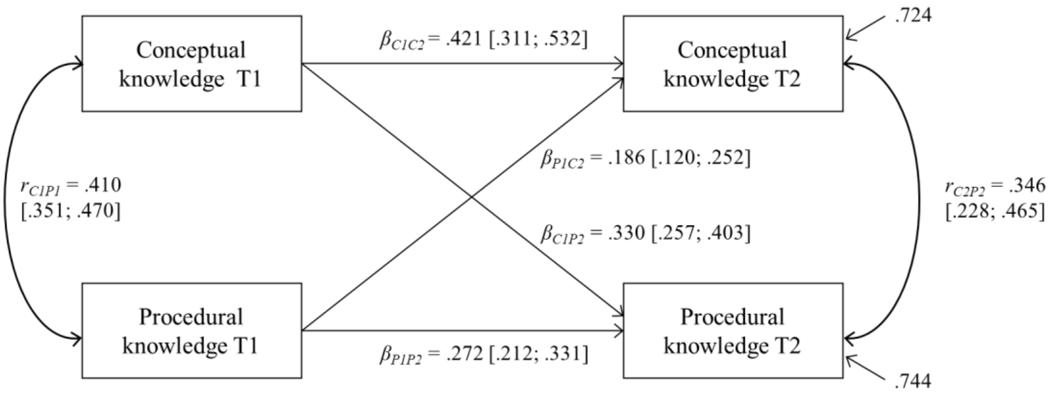
Cross-lagged Panel Model of Moderator Level Kindergarten / Preschool



Note. Cross-lagged panel model of conceptual knowledge at measurement point one (C1), procedural knowledge at measurement point one (P1), conceptual knowledge at measurement point two (C2), procedural knowledge at measurement point two (P2) as well as the cross-sectional (r_{C1P1} and r_{C2P}), auto-regressive (β_{C1C2} and β_{P1P2}) and cross-lagged paths between the four variables (β_{C1P2} and β_{P1C2}). The Hessian of this model was not convex - results must be interpreted with caution; 95% Confidence intervals estimated by setting variances without SEs at 0; k = 4 to 9 for the six bivariate correlations.

Figure 11

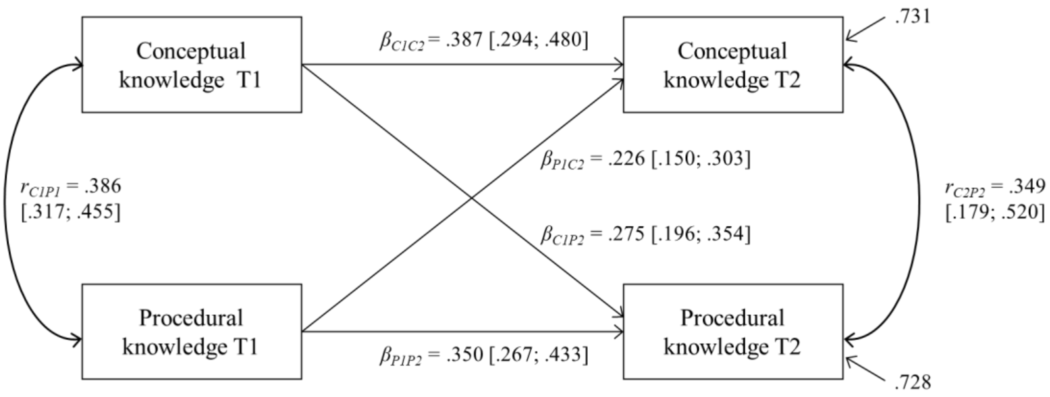
Cross-lagged Panel Model of Moderator Level Primary School



Note. Cross-lagged panel model of conceptual knowledge at measurement point one (C1), procedural knowledge at measurement point one (P1), conceptual knowledge at measurement point two (C2), procedural knowledge at measurement point two (P2) as well as the cross-sectional (r_{C1P1} and r_{C2P}), auto-regressive (β_{C1C2} and β_{P1P2}) and cross-lagged paths between the four variables (β_{C1P2} and β_{P1C2}); k = 14 to 18 for the six bivariate correlations.

Figure 12

Cross-Lagged Panel Model of Moderator Level Secondary School



Note. Cross-lagged panel model of conceptual knowledge at measurement point one (C1), procedural knowledge at measurement point one (P1), conceptual knowledge at measurement point two (C2), procedural knowledge at measurement point two (P2) as well as the cross-sectional (r_{C1P1} and r_{C2P}), auto-regressive (β_{C1C2} and β_{P1P2}) and cross-lagged paths between the four variables (β_{C1P2} and β_{P1C2}); k = 17 to 19 for the six bivariate correlations.

Table 4

Moderator Analysis School Level

Level	Comparison level	*k*	Change in β_{C1P2}	95% CI	Change in β_{P1C2}	95% CI
Kindergarten / Preschool	Primary school	14 - 18	.228	-	-.130	-
	Secondary school	23 - 27	.076	-	.019	-
Primary school	Secondary school	33 - 35	-.054	[-.160; .053]	.048	[-.056; .152]

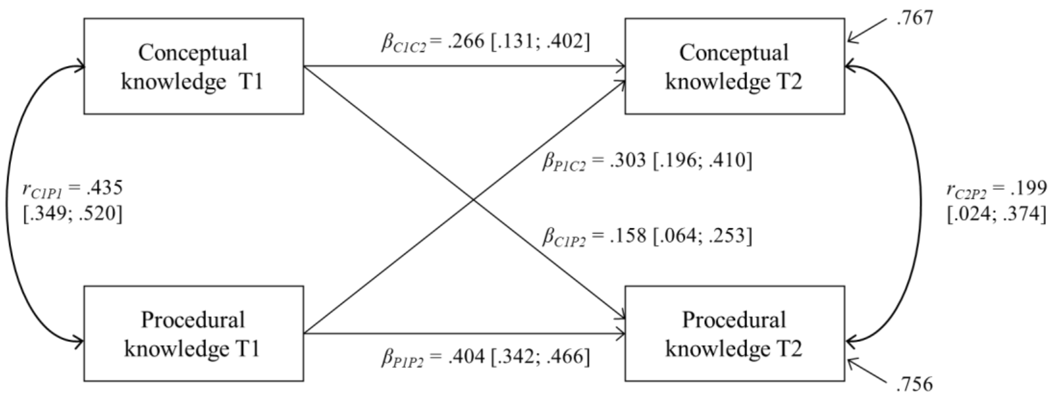
Note. Overview of moderator analysis of categorical moderator *School Level*. Level - Moderator level, Comparison level - Moderator level that is compared to the level in the column Level, *k* - Number of effect sizes for each bivariate correlation, Change in β_{C1P2} - Change in path coefficient β_{C1P2} from Level to Comparison Level; 95% CI - 95% Confidence interval, Change in $\boldsymbol{\beta_{P1C2}}$ - Change in path coefficient β_{P1C2} from Level to Comparison Level. Missings are indicated as -. These confidence Intervals could not be estimated due to convergence problems.

5.5.4.3 Subdomain Mathematics. We compared the moderator levels of *Subdomain Mathematics* to test hypothesis H4; the predictive relations between conceptual and procedural knowledge differ between the subdomains of mathematics. This categorical moderator initially included five levels. The levels *Other* and *Several* were left out of the analyses because they only contained three and four effect sizes, respectively. Therefore, we inspected three levels: *Counting and Whole-Number Arithmetic*, *Rational Numbers*, and *Algebra*. The standardized path coefficients of the three included moderator levels are visualized in Figures 13 to 15. The model on *Rational Numbers* must be interpreted with caution because the Hessian was not convex and the information matrix was not positive definite.

The predictive relations between conceptual and procedural knowledge differed descriptively in strength and symmetry between the levels: The predictive relation from procedural to conceptual knowledge (β_{P1C2}) was more substantial in the subdomain *Counting and Whole-Number Arithmetic*. In contrast, the predictive relation from conceptual to procedural knowledge (β_{C1P2}) was stronger in the subdomain of *Rational Numbers*. In the subdomain of *Algebra*, both relations (β_{C1P2} and β_{P1C2}) were equally strong.

Figure 13

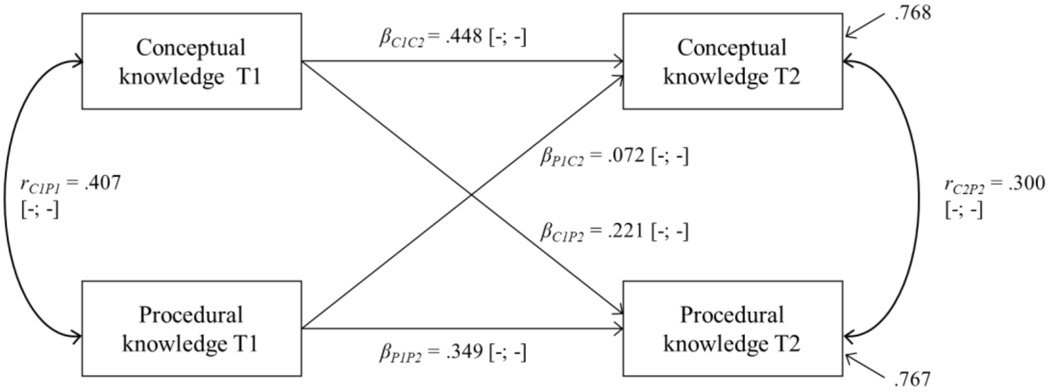
Cross-Lagged Panel Model of Moderator Level Counting and Whole-Number Arithmetic



Note. Cross-lagged panel model of conceptual knowledge at measurement point one (C1), procedural knowledge at measurement point one (P1), conceptual knowledge at measurement point two (C2), procedural knowledge at measurement point two (P2) as well as the cross-sectional (r_{C1P1} and r_{C2P}), auto-regressive (β_{C1C2} and β_{P1P2}) and cross-lagged paths between the four variables (β_{C1P2} and β_{P1C2}); k = 8 to 14 for the six bivariate correlations.

Figure 14

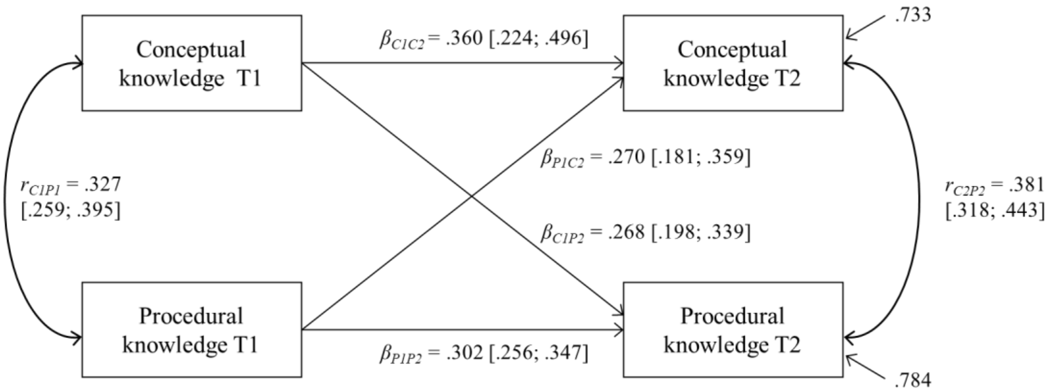
Cross-Lagged Panel Model of Moderator Level Rational Numbers



Note. Cross-lagged panel model of conceptual knowledge at measurement point one (C1), procedural knowledge at measurement point one (P1), conceptual knowledge at measurement point two (C2), procedural knowledge at measurement point two (P2) as well as the cross-sectional (r_{C1P1} and r_{C2P}), auto-regressive (β_{C1C2} and β_{P1P2}) and cross-lagged paths between the four variables (β_{C1P2} and β_{P1C2}). The Hessian of this model was not convex - results must be interpreted with caution. Missings are indicated as -. Confidence intervals could not be estimated due to convergence problems; k = 11 for the six bivariate correlations.

Figure 15

Cross-Lagged Panel Model of Moderator Level Algebra



Note. Cross-lagged panel model of conceptual knowledge at measurement point one (C1), procedural knowledge at measurement point one (P1), conceptual knowledge at measurement point two (C2), procedural knowledge at measurement point two (P2) as well as the cross-sectional (r_{C1P1} and r_{C2P}), auto-regressive (β_{C1C2} and β_{P1P2}) and cross-lagged paths between the four variables (β_{C1P2} and β_{P1C2}); k = 15 to 19 for the six bivariate correlations.

The levels *Counting and Whole-Number Arithmetic*, *Rational Numbers*, and *Algebra* were compared in pairs in three separate analyses. Again, the Hessian was not convex, and the information matrix was not positive definitive for the moderator model of *Rational Numbers*. These results must be interpreted with caution. There were no significant differences in the predictive relations between conceptual and procedural knowledge in each comparison of the three levels (see Table 5). Concluding the analyses, hypothesis H4 was not confirmed.

Table 5

Moderator Analysis Subdomain of Mathematics

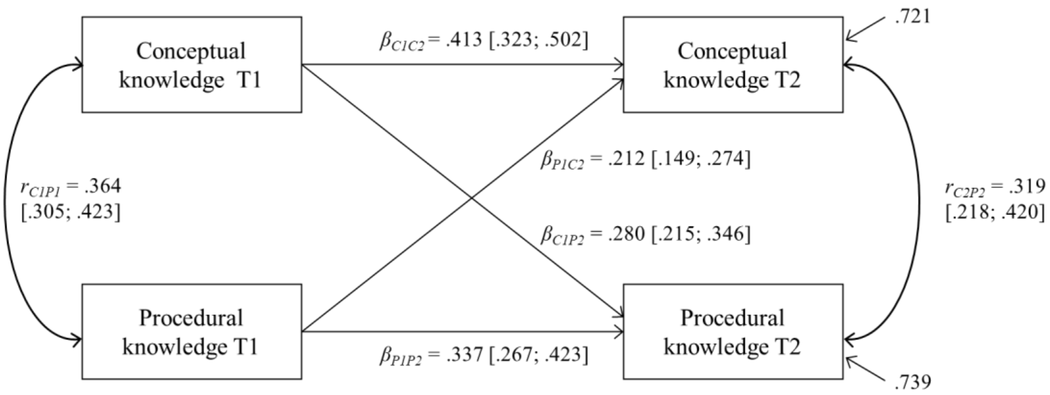
Level	Comparison Level	*k*	Change in β_{C1P2}	95% CI	Change in β_{P1C2}	95% CI
Counting and Whole-Number Arithmetic	Rational Numbers	19 - 25	.120	-	-.244	-
	Algebra	25 - 31	.138	-	-.092	-
Rational Numbers	Algebra	27 - 30	-.003	-	.151	-

Note. Overview of moderator analysis of categorical moderator *Subdomain of Mathematics*. Level - Moderator level, Comparison Level - Moderator level that is compared to the level in the column Level, *k* - Number of effect sizes for each bivariate correlation, Change in β_{C1P2} - Change in path coefficient β_{C1P2} from Level to Comparison Level; 95% CI - 95% Confidence interval, Change in $\boldsymbol{\beta_{P1C2}}$ - Change in path coefficient β_{P1C2} from Level to Comparison Level. Missings are indicated as -. These confidence intervals could not be estimated due to convergence problems.

5.5.4.4 Intervention. The categorical moderator included two levels, *Intervention* and *No Intervention*. We analyzed this moderator to test hypothesis H5; the predictive relations between conceptual and procedural knowledge are stronger in studies in which an intervention took place than in those without an intervention. The results of the level *No Intervention* model must be interpreted with caution; the Hessian was not convex, and the information matrix was not positive definite. For both models, the path coefficients from conceptual to procedural knowledge (β_{C1P2}) were slightly more substantial than those from procedural to conceptual knowledge (β_{P1C2}, see Figures 16 and 17).

Figure 16

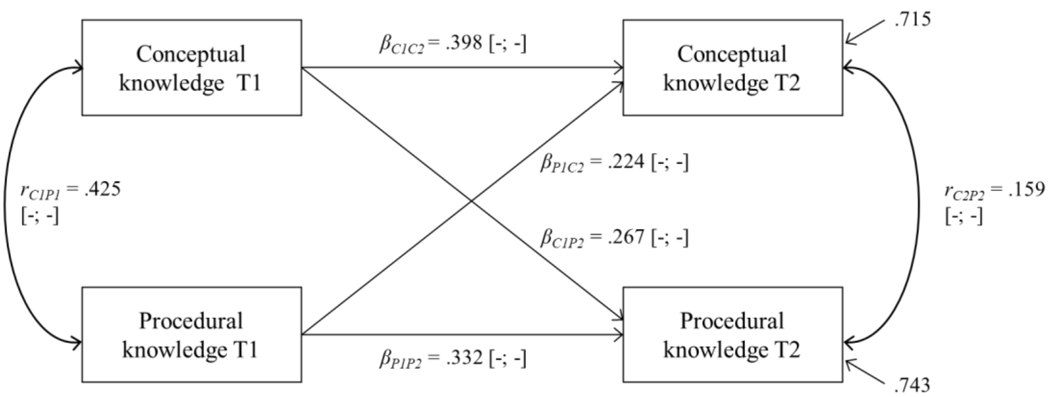
Cross-Lagged Panel Model of Moderator Level Intervention



Note. Cross-lagged panel model of conceptual knowledge at measurement point one (C1), procedural knowledge at measurement point one (P1), conceptual knowledge at measurement point two (C2), procedural knowledge at measurement point two (P2) as well as the cross-sectional (r_{C1P1} and r_{C2P}), auto-regressive (β_{C1C2} and β_{P1P2}) and cross-lagged paths between the four variables (β_{C1P2} and β_{P1C2}); k = 25 to 27 for the six bivariate correlations.

Figure 17

Cross-Lagged Panel Model of Moderator Level No Intervention



Note. Cross-lagged panel model of conceptual knowledge at measurement point one (C1), procedural knowledge at measurement point one (P1), conceptual knowledge at measurement point two (C2), procedural knowledge at measurement point two (P2) as well as the cross-sectional (r_{C1P1} and r_{C2P}), auto-regressive (β_{C1C2} and β_{P1P2}) and cross-lagged paths between the four variables (β_{C1P2} and β_{P1C2}). The Hessian of this model was not convex - results must be interpreted with caution. Missings are indicated as -. Confidence intervals could not be estimated due to convergence problems; k = 14 to 22 for the six bivariate correlations.

To test for moderating influences, we compared the two levels of the moderator. The results of this analysis should be interpreted with caution because, again, the Hessian was not convex and the information matrix was not positive definite. For the studies with an intervention, the predictive relations from conceptual to procedural knowledge changed by β_{C1P2} = .046 (95% CI [-; -]) compared to studies without an intervention. The relations from procedural to conceptual knowledge decreased by β_{P1C2} = -.082 (95% CI [-; -]) in this comparison. There were no significant differences between the models with and without intervention, and we rejected hypothesis H5.

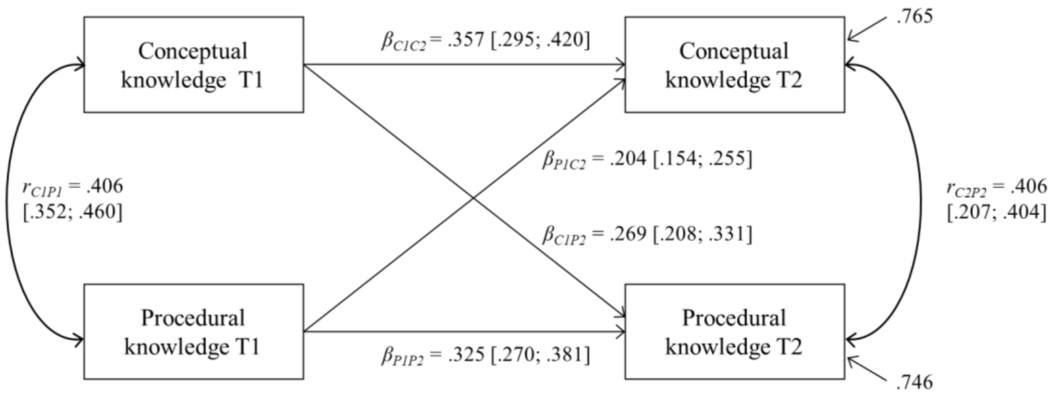
5.5.5 Exploratory Analyses

5.5.5.1 Time Lag. The time between the measurements at T1 and T2 might moderate the predictive relations. To answer this, we analyzed the continuous moderator *Time Lag*. The Hessian was not convex and the information matrix was not positive definite for the moderation model. Therefore, the results must be interpreted with caution. According to the model, no moderating effect existed on the predictive relations between conceptual and procedural knowledge. The path coefficients did not change significantly when the time lag between measurements increased (β_{C1P2} = -.009, 95% CI [-; -]; β_{P1C2} = .049, 95% CI [-; -]).

5.5.5.2 Broad Task Type of Conceptual Knowledge. The moderator *Broad Task Type of Conceptual Knowledge* included three levels. These were *Implicit Task Type*, *Explicit Task Type*, and *Both*. The results of the *Explicit Task Type* and *Both* models must be interpreted cautiously because the Hessian was not convex and the information matrix was not positive definite for these models. For all three levels, the predictive relation from conceptual to procedural knowledge (β_{C1P2}) was descriptively stronger than vice versa (β_{P1C2}, see Figure 18 to 20).

Figure 18

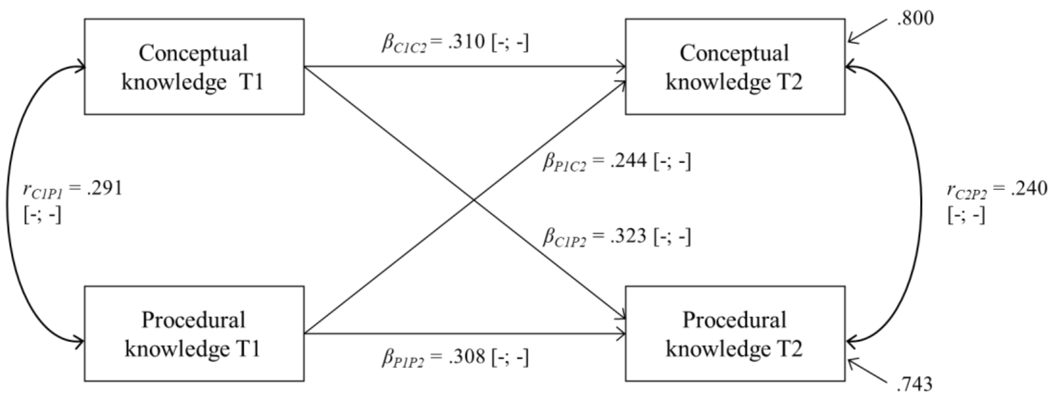
Cross-Lagged Panel Model of Moderator Level Implicit Task Type



Note. Cross-lagged panel model of conceptual knowledge at measurement point one (C1), procedural knowledge at measurement point one (P1), conceptual knowledge at measurement point two (C2), procedural knowledge at measurement point two (P2) as well as the cross-sectional (r_{C1P1} and r_{C2P}), auto-regressive (β_{C1C2} and β_{P1P2}) and cross-lagged paths between the four variables (β_{C1P2} and β_{P1C2}); k = 25 to 32 for the six bivariate correlations.

Figure 19

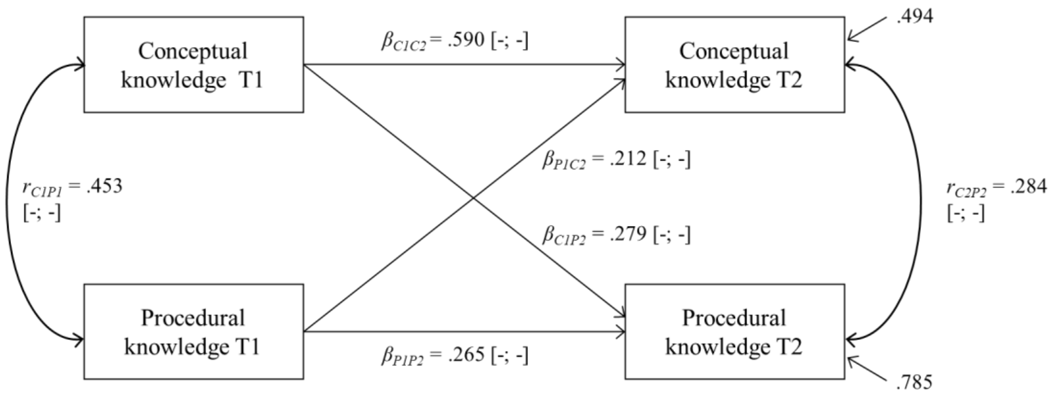
Cross-Lagged Panel Model of Moderator Level Explicit Task Type



Note. Cross-lagged panel model of conceptual knowledge at measurement point one (C1), procedural knowledge at measurement point one (P1), conceptual knowledge at measurement point two (C2), procedural knowledge at measurement point two (P2) as well as the cross-sectional (r_{C1P1} and r_{C2P}), auto-regressive (β_{C1C2} and β_{P1P2}) and cross-lagged paths between the four variables (β_{C1P2} and β_{P1C2}). The Hessian of this model was not convex - results must be interpreted with caution. Missings are indicated as -. Confidence intervals could not be estimated due to convergence problems; k = 7 for the six bivariate correlations.

Figure 20

Cross-Lagged Panel Model of Moderator Level Both Task Types



Note. Cross-lagged panel model of conceptual knowledge at measurement point one (C1), procedural knowledge at measurement point one (P1), conceptual knowledge at measurement point two (C2), procedural knowledge at measurement point two (P2) as well as the cross-sectional (r_{C1P1} and r_{C2P}), auto-regressive (β_{C1C2} and β_{P1P2}) and cross-lagged paths between the four variables (β_{C1P2} and β_{P1C2}). The Hessian of this model was not convex - results must be interpreted with caution. Missings are indicated as -. Confidence intervals could not be estimated due to convergence problems; k = 6 to 9 for the six bivariate correlations.

We compared the three moderator levels in separate pairwise analyses regarding the predictive relations between conceptual and procedural knowledge. None of the levels differed significantly in the cross-lagged path coefficients β_{C1P2} and β_{P1C2}. The results of the comparison of *Implicit Task Type* and *Both* must be interpreted with caution because the Hessian of the model was not convex and the information matrix was not positive definite. The three inspected moderator levels had no moderating influence on the predictive relations between conceptual and procedural knowledge (see Table 6).

5.5.5.3 Specific Task Type of Conceptual Knowledge. The moderator *Specific Task Type of Conceptual Knowledge* initially included 13 levels. Four of these 13 levels were found in the data: *Translate Quantities Between Representational Systems*, *Compare Quantities*, *Evaluate Unfamiliar Procedures*, *Explain Judgments*, and *Several*. The results of the models

Table 6

Moderator Analysis Broad Task Type of Conceptual Knowledge

Level	Comparison Level	k	Change in β_{C1P2}	95% CI	Change in β_{P1C2}	95% CI
Implicit	Explicit	32 – 39	.052	[-.071; .174]	.035	[-.066; .135]
	Both	32 – 40	.030	[-.119; .178]	.019	[-.113; .150]
Implicit	Both	13 – 16	.156	-	.110	-

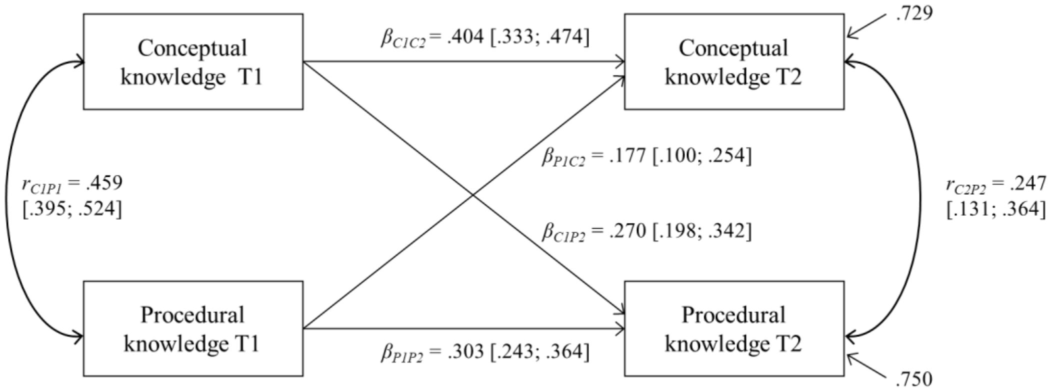
Note. Overview of moderator analysis of categorical moderator *Broad Task Type of Conceptual Knowledge*. Level - Moderator level, Comparison Level - Moderator level that is compared to the level in the column Level, k - Number of effect sizes for each bivariate correlation, Change in β_{C1P2} - Change in path coefficient β_{C1P2} from Level to Comparison Level; 95% CI - 95% Confidence interval, Change in $\boldsymbol{\beta_{P1C2}}$ - Change in path coefficient β_{P1C2} from Level to Comparison Level. Missings are indicated as -. These confidence intervals could not be estimated due to convergence problems.

Evaluate Unfamiliar Procedures and *Explain Judgments* must be interpreted with caution because the Hessian of both models was not convex and the information matrix was not positive definite. The predictive relation from conceptual to procedural knowledge (β_{C1P2}) was descriptively stronger than the relation from procedural to conceptual knowledge (β_{P1C2}) in all inspected models (see Figures 21 to 25).

None of the levels differed in their predictive relations between conceptual and procedural knowledge compared to the other levels (see Table 7). Therefore, the specific task type of conceptual knowledge did not moderate the predictive relations between conceptual and procedural knowledge.

Figure 21

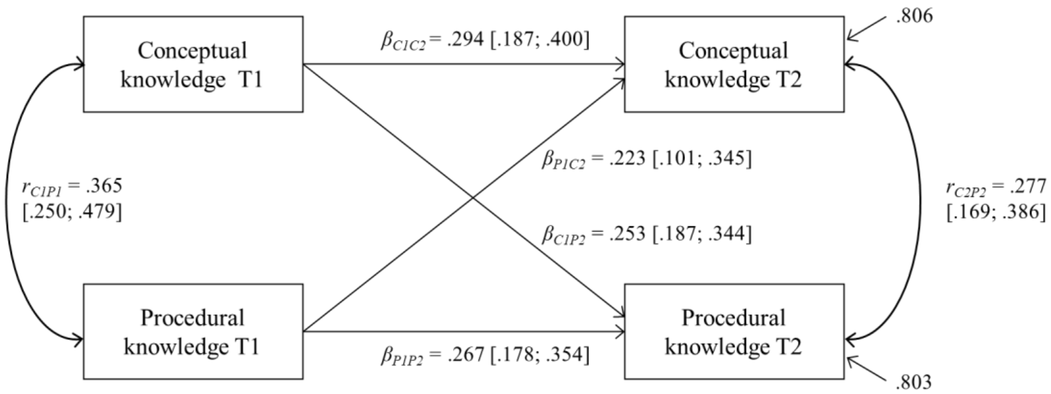
Cross-Lagged Panel Model of Moderator Level Translate Quantities Between Representational Systems



Note. Cross-lagged panel model of conceptual knowledge at measurement point one (C1), procedural knowledge at measurement point one (P1), conceptual knowledge at measurement point two (C2), procedural knowledge at measurement point two (P2) as well as the cross-sectional (r_{C1P1} and r_{C2P}), auto-regressive (β_{C1C2} and β_{P1P2}) and cross-lagged paths between the four variables (β_{C1P2} and β_{P1C2}); k = 15 to 24 for the six bivariate correlations.

Figure 22

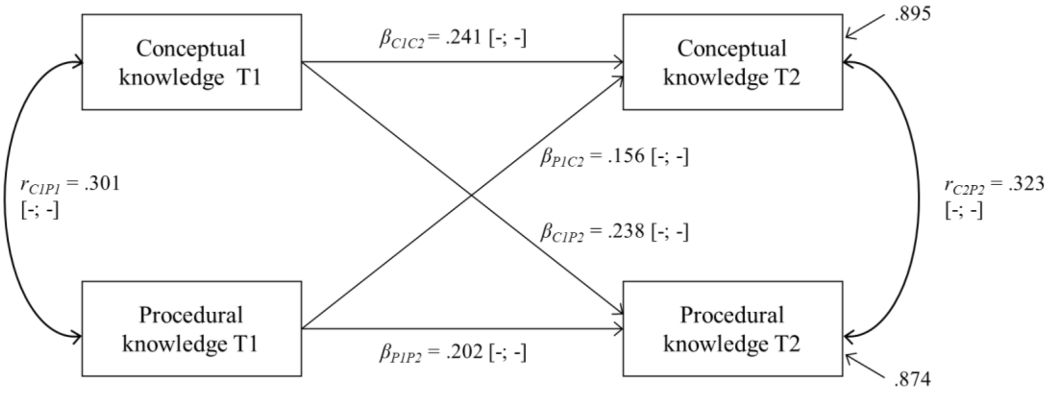
Cross-Lagged Panel Model of Moderator Level Compare Quantities



Note. Cross-lagged panel model of conceptual knowledge at measurement point one (C1), procedural knowledge at measurement point one (P1), conceptual knowledge at measurement point two (C2), procedural knowledge at measurement point two (P2) as well as the cross-sectional (r_{C1P1} and r_{C2P}), auto-regressive (β_{C1C2} and β_{P1P2}) and cross-lagged paths between the four variables (β_{C1P2} and β_{P1C2}); k = 6 to 9 for the six bivariate correlations.

Figure 23

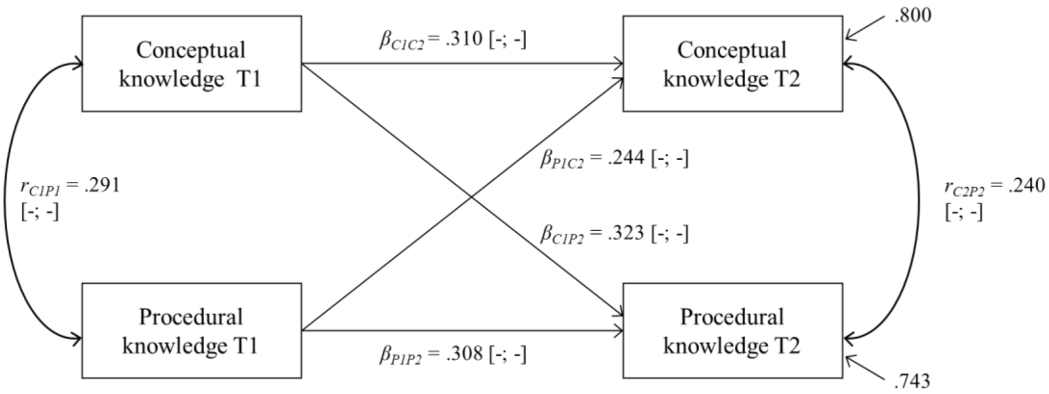
Cross-Lagged Panel Model of Moderator Level Evaluate Unfamiliar Procedures



Note. Cross-lagged panel model of conceptual knowledge at measurement point one (C1), procedural knowledge at measurement point one (P1), conceptual knowledge at measurement point two (C2), procedural knowledge at measurement point two (P2) as well as the cross-sectional (r_{C1P1} and r_{C2P}), auto-regressive (β_{C1C2} and β_{P1P2}) and cross-lagged paths between the four variables (β_{C1P2} and β_{P1C2}). The Hessian of this model was not convex - results must be interpreted with caution; Missings are indicated as -. Confidence intervals could not be estimated due to convergence problems; k = 4 to 7 for the six bivariate correlations.

Figure 24

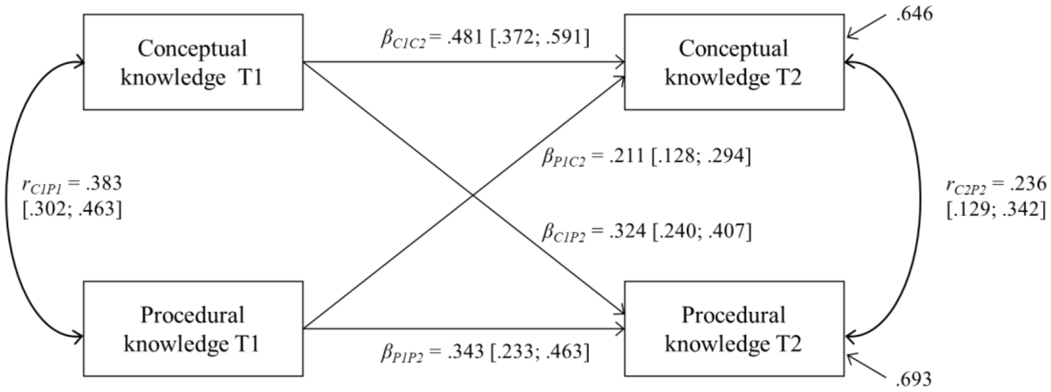
Cross-Lagged Panel Model of Moderator Level Explain Judgments



Note. Cross-lagged panel model of conceptual knowledge at measurement point one (C1), procedural knowledge at measurement point one (P1), conceptual knowledge at measurement point two (C2), procedural knowledge at measurement point two (P2) as well as the cross-sectional (r_{C1P1} and r_{C2P}), auto-regressive (β_{C1C2} and β_{P1P2}) and cross-lagged paths between the four variables (β_{C1P2} and β_{P1C2}). The Hessian of this model was not convex - results must be interpreted with caution; Missings are indicated as -; Confidence intervals could not be estimated due to convergence problems; k = 7 for the six bivariate correlations.

Figure 25

Cross-Lagged Panel Model of Moderator Level Several Tasks



Note. Cross-lagged panel model of conceptual knowledge at measurement point one (C1), procedural knowledge at measurement point one (P1), conceptual knowledge at measurement point two (C2), procedural knowledge at measurement point two (P2) as well as the cross-sectional (r_{C1P1} and r_{C2P}), auto-regressive (β_{C1C2} and β_{P1P2}) and cross-lagged paths between the four variables (β_{C1P2} and β_{P1C2}); k = 12 to 16 for the six bivariate correlations.

Table 7

Moderator Analysis Specific Task Type of Conceptual Knowledge

Level	Comp. Level	*k*	Change in β_{C1P2}	95% CI	Change in β_{P1C2}	95% CI
Translate Quantities	Compare Quantities	22 - 33	-.021	[-.125; .103]	.046	[-.078; .170]
	Evaluate Unfamiliar Procedures	22 - 28	-.044	[-.190; .102]	-.003	[-.174; .107]
	Explain Judgments	22 - 31	.054	[-.067; .175]	.064	[-.052; .180]
	Several	30 - 37	.061	-	.058	-
Compare Quantities	Evaluate Unfamiliar Procedures	11 - 16	-.031	[-.185; .123]	-.088	[-.271; .096]
	Explain Judgments	14 - 16	.065	[-.065; .195]	.024	[-.123; .171]
	Several	19 - 25	.060	[-.064; .184]	-.003	[-.145; .139]
Evaluate Unfamiliar Procedures	Explain Judgments	11 - 14	.078	-	.074	-
	Several	16 - 23	.088	[-.056; .232]	.046	[-.095; .187]
Explain Judgments	Several	19 - 23	.002	[-.127; .130]	-.029	[-.140; .083]

Note. Overview of moderator analysis of categorical moderator *Specific Task Type of Conceptual Knowledge*. Level - Moderator level, Comp. Level - Moderator level that is compared to the level in the column Level, *k* - Number of effect sizes for each bivariate correlation, Change in β_{C1P2} - Change in path coefficient β_{C1P2} from Level to Comparison Level; 95% CI - 95% Confidence interval, Change in $\boldsymbol{\beta_{P1C2}}$ - Change in path coefficient β_{P1C2} from Level to Comparison Level. Missings are indicated as -. These confidence intervals could not be estimated due to convergence problems.

5.6 Discussion

5.6.1 Main Findings

This meta-analysis examined the predictive relations between conceptual and procedural knowledge in mathematics as path coefficients in a cross-lagged panel model. The model was used to answer the questions of how strong the predictive relations between the two knowledge types are and whether these relations are bidirectional and symmetrical. Bidirectional predictive relations indicate that conceptual and procedural knowledge develop iteratively (Rittle-Johnson, 2017).

Conceptual knowledge predicted changes in procedural knowledge ($\beta_{C1P2} = .270$) and simultaneously, procedural knowledge also predicted changes in conceptual knowledge ($\beta_{P1C2} = .220$). These two relations were similarly strong with large cross-lagged effects (Orth et al., 2022). In addition, they were bidirectional and almost symmetrical. The main hypothesis H1 was confirmed. The results support the Iterative Model with gradual bidirectional improvements in both knowledge types (Rittle-Johnson et al., 2001). Moreover, these significant predictive relations between knowledge types over time are consistent with the predictive relations found in the cross-lagged panel model by Schneider and colleagues (2011), which were also bidirectional and symmetrical.

There are several possible reasons why the predictive relations in the model seem relatively weak compared to previous findings (e.g., Hecht & Vagi, 2010). First, compared to bivariate correlations, cross-lagged effects are controlled for prior levels of the predicted variable and concurrent correlations at measurement point one (Orth et al., 2022). Therefore, cross-lagged effects are generally smaller than effects of bivariate correlations. Another reason might be that in the overall model, characteristics of the sample and subdomains balanced out the predictive relations: While younger children depend more on their procedural knowledge (e.g., Frye et al., 1989; LeFevre et al., 2006), conceptual knowledge seems more important for older learners (Gelman & Williams, 1998; Halford, 1993; Hecht et al., 2003; Hiebert &

Lefevre, 1986; Schneider et al., 2009). Also, the strength of the predictive relations might differ between subdomains of mathematics (e.g., Rittle-Johnson & Sielger, 1998).

Findings from our analysis make a valuable contribution on both a theoretical and practical level. Theoretically, the results help to understand developmental patterns of conceptual and procedural knowledge providing valuable implications on the validity of existing viewpoints on the relationship. On a practical level, the results help educators to decide the order of teaching conceptual or procedural knowledge and to adjust it due to naturally occurring developmental patterns of the student's knowledge acquisition.

5.6.2 Moderator Variables

The predictive relations were stable across all four moderators. We investigated to what extent the learners' age, school level, the subdomain of mathematics, and the presence of an intervention moderated the predictive relations between conceptual and procedural knowledge.

The mean age of the learners did not influence the predictive relations between conceptual and procedural knowledge. Hypothesis H2 must be rejected. One explanation for this finding is that the differences in the predictive relations in the data were not linear. This was inspected by plotting the effect sizes in dependence on the mean age. Therefore, they cannot be portrayed by the learners' age as a continuous moderator.

The predictive relations between conceptual and procedural knowledge did not differ depending on the school level. Of the six levels, only three (kindergarten/preschool, primary school, and secondary school) could be compared. There was no statistically significant difference in the relation from procedural knowledge to conceptual knowledge and vice versa between kindergarten/preschool, primary, and secondary school. Hence, hypothesis H3 was not confirmed for these three school levels. The results of the comparison between kindergarten/preschool and each of the two school levels must be interpreted with caution because the moderation models did not converge successfully. It is conceivable that the predictive relations between the two knowledge types differ less between the levels after starting school than between kindergarten and school. Regular school instructions might support conceptual as well as procedural knowledge. These are contained in all school levels and reduce the differences in the predictive relations between students. The descriptive results of the comparison models are consistent with these assertions. Children in kindergarten or preschool descriptively showed stronger influences of procedural on conceptual knowledge, whereas for children in primary and secondary school, both relations were more similar.

The subdomains of mathematics did not influence the predictive relations between conceptual and procedural knowledge. Previous findings purported differences in the predictive relations between subdomains (e.g., Rittle-Johnson & Siegler, 1998). In the domain of rational numbers, conceptual knowledge was found to have a stronger influence on procedural knowledge (Byrnes & Wasik, 1991). On the other hand, findings in the subdomain of algebra supported the greater impact of procedural knowledge on subsequent conceptual knowledge (Lauritzen, 2012). We did not find statistically significant differences between the subdomains of mathematics, and hypothesis H4 must be rejected. However, the descriptive results of our comparison models again align with the assumptions of our hypothesis H4. The model on counting and whole-number arithmetic included stronger influences of procedural on conceptual knowledge. The opposite was present for the model of rational numbers. Last, the model of algebra showed a symmetrical impact on the knowledge types. At least two possible conclusions can be drawn from this result. First, contrary to our descriptive results, the predictive relations are mostly stable across the subdomains of mathematics and, therefore, did not differ significantly in the analysis. This assumption implies that the previous findings (e.g., Byrnes & Wasik, 1991; Lauritzen, 2012) only partly displayed the relations between knowledge types. The other assumption contains the restriction that the categories in this meta-analysis were too broad. Consequently, the predictive relations differed in parts of the subdomains (e.g., Rittle-Johnson & Siegler, 1998) within the categories and compensated for possible differences.

Whether an intervention was carried out did not influence the strength of both predictive relations. We hypothesized that the predictive relations between conceptual and procedural knowledge would be stronger for the studies in which an intervention was carried out. Evidence shows that interventions can increase knowledge types (Canobi, 2009) and their relations (Blöte et al., 2001; Hiebert & Wearne, 1996). This analysis did not confirm this result, and hypothesis H5 must also be rejected. An alternate explanation for this finding could be that the interventions in the included studies differed in their goals and methods and that possible effects of interventions counterbalanced each other.

The predictive relations between conceptual and procedural knowledge were stable across the exploratory moderators. We conducted exploratory analyses on the time lag between measurements and the broad as well as specific task types of the assessment of conceptual knowledge. None of these moderators showed influences on the predictive relations between conceptual and procedural knowledge.

5.6.3 Limitations

Some methodological aspects of this meta-analysis restrict the generalizability of the results. First, there is the question about the validity of the measurements of conceptual and procedural knowledge. There are no widely used standardized approaches up to this point. Recent findings show that conceptual and procedural fraction knowledge are empirically separable when using appropriate measures (Lenz et al., 2020). However, most included studies did not use such exclusively developed instruments. This makes it difficult to attribute test results to only one of the two knowledge types (Rittle-Johnson & Schneider, 2014). For example, when participants solve given routine problems, both knowledge types are involved in the process (Braithwaite & Sprague, 2021). This meta-analysis's cross-sectional correlations between conceptual and procedural knowledge at T1 and T2 are medium-sized. Consequently, the measurements' inaccuracy regarding the knowledge types' assessments may have only slightly overestimated the results.

Second, the aggregation of the dependent effect sizes by averaging them on the sample level yields some disadvantages: Averaging effect sizes ignores possible within-study variability (López-López et al., 2018). This loss of information limits the possibility of examining characteristics to evaluate effect size variability (Wilson et al., 2016). Also, the resulting number of independent effect sizes was relatively small but sufficiently high to conclude that the OSMASEM works well under the assumption of missing data at random (Jak & Cheung, 2020).

Third, effect sizes could be missing at random or systematically (Jak & Cheung, 2020). The probability of effect sizes missing systematically is reduced by the inclusion of peer-reviewed articles as well as grey literature in which the bivariate correlations on conceptual and procedural knowledge were not the main aspects of the research. Also, the funnel plots and Egger's regressions did not indicate the presence of a publication bias. Therefore, this limitation might only slightly influence the main results.

The last limitation concerns the decision to use a CLPM. Several model assumptions can be criticized in applied contexts (Hamaker et al., 2015; Lucas, 2023; Mund & Nestler, 2019). However, most of the included primary studies contained either CLPM or were studies with only two waves of data. In this case, using a CLPM had more advantages than other advanced approaches, which would have needed a different database with more waves of data.

5.6.4 Implications

5.6.4.1 Implications for Mathematics Education. The finding of bidirectional predictive relations between conceptual and procedural knowledge yields several implications for mathematics education. Researchers, as well as practitioners, can use the results of this meta-analysis to improve mathematics education.

Starting in kindergarten or preschool, a solid knowledge base can be promoted to enable subsequent improvements in both knowledge types. Building on the bidirectional and symmetrical predictive relations we found, we agree that there might not be an optimal instruction ordering (Rittle-Johnson et al., 2015). Adding to this, mathematics education can better support children in their learning process at school by acknowledging that they differ in their initial knowledge (Rittle-Johnson et al., 2001) and in their knowledge profiles (Hecht & Vagi, 2012). Educators should avoid a generally predetermined order of knowledge acquisition in school syllabuses. This allows more flexibility for schools and teachers to adjust the curriculum to the students' individual needs.

Furthermore, although the moderator analyses did not show significant differences between the moderator levels, it might be beneficial for teachers and instructors to pay attention to the conditions of the learners and content. Previous research concluded a stronger influence of procedural knowledge on young children's conceptual knowledge (e.g., LeFevre et al., 2006). However, for older learners, conceptual knowledge seemed more important (e.g., Schneider et al., 2009). We descriptively confirmed these assumptions based on school levels. Furthermore, relations were previously found to differ between subdomains (e.g., Rittle-Johnson & Siegler, 1998). Again, our descriptive results pointed in the same direction. Concluding, it might be conducive to success in mathematics education to adjust the focus of instructions to relevant aspects of content and learners besides their knowledge profiles. With this, differences between subgroups in the predictive relations can be used to promote both knowledge types best.

5.6.4.2 Future Studies. This meta-analysis showed that there are bidirectional predictive relations between conceptual and procedural knowledge. These relations remained stable across all examined moderators. Future studies are essential to investigate the possible influences of other variables in other knowledge domains.

We limited the analysis to the knowledge domain of mathematics because of missing primary studies in other STEM domains. We excluded studies that measured conceptual and procedural knowledge with time limits (e.g., Bailey et al., 2012) or used solution time as an outcome (e.g., Paul et al., 2019). These decisions limit the generalizability of the results, but they also improve the congruity of the analysis by providing similar definitions and measurements. Future studies should focus on other knowledge domains to answer whether conceptual and procedural knowledge can be differentiated in these; and, if they are, how they are interrelated.

Moreover, there might be moderating influences that this meta-analysis did not capture. On the one hand, some of the examined moderators could be inspected more closely. For example, whether an intervention was carried out was examined dichotomously. Instead, a distinction between types and goals of interventions might give further insights. On the other hand, including other moderators would be essential. The regular school instruction itself might influence the relations. Instructions could differ according to the national guidelines (e.g., NCTM, 2014) in starting with either conceptual or procedural knowledge. This moderator would affect the initial learners' knowledge base and yield further insights into their knowledge profile development.

The OSMASEM, an approach to fit correlation matrices to a structural equation model in meta-analyses, is still under development. Some analyses cannot be conducted in this model up to this point. Examples are corrections for dependent effect sizes or moderator analyses with categorical moderators with more than two levels. Future meta-analyses might help reinvestigate the predictive relations when these features are available.

Adding to this, it could be valuable to consider other models than a CLPM to examine the predictive relations between conceptual and procedural knowledge in the future. Advanced models have some benefits compared to the CLPM. One is that these models differentiate between-person and within-person variance. This can help to interpret cross-lagged effects, but it also contains other possible problems (Orth et al., 2021). However, advanced models like the Random-Intercept Cross-Lagged Panel Model have the prerequisite that at least three measurement points are available in primary studies (Mund & Nestler, 2019). Future studies in this research area should include at least three measurement points to provide appropriate data.

Last, most previous studies focused on correlational data and less on longitudinal data. Experiments are even more scarcely found. More randomized controlled trials would be necessary to examine possible causal effects between conceptual and procedural knowledge. Systematic variation of the amount of conceptual and procedural knowledge at T1 enables more insight into the predictive relations depending on the learner's initial knowledge. In addition, instructional interventions could be systematically employed to foster predictive relations.

5.6.5 Conclusion

This meta-analysis showed positive bidirectional and almost symmetrical predictive relations between conceptual and procedural knowledge in the domain of mathematics. At least four different theoretical viewpoints on the predictive relations were purported by previous research. Our finding aligns with the main statement of the Iterative Model (Rittle-Johnson et al., 2001). The standardized path coefficients for the predictive relations in our cross-lagged panel model were large and significant. The predictive relations were stable across the inspected moderators. To better understand the possible impact of moderators, future studies should be carried out using the developing advantages of the OSMASEM (Jak et al., 2021). Concluding, this meta-analysis confirmed the theoretical view of the Iterative Model on the predictive relations between conceptual and procedural knowledge in the domain of mathematics.

6 Study 2: How Strongly Do Motivational Constructs Mediate the Influence of Prior Knowledge on Posttest Knowledge? A Meta-Analytic Investigation of Moderated Mediation Effects

6.1 Abstract

A recent meta-analysis (Simonsmeier et al., 2022) averaged over 8000 effect sizes indicating the effect of prior knowledge on learning. They found that the stability of individual differences in knowledge from before to after learning was high ($r_P^+ = .534$), even though the predictive power of prior knowledge for knowledge gains was low ($r_{NG}^+ = -.059$). This raises the question of why individual differences in knowledge are stable over time, a finding that remained even after controlling for intelligence. One possible explanation is that prior knowledge might improve motivation and related constructs leading to increased new knowledge. We conducted a meta-analysis to test this hypothesis. We examined how strongly motivational constructs (interest, self-concept, self-efficacy, intrinsic motivation, extrinsic motivation) mediate the effect of prior knowledge on knowledge after learning and what other variables (e.g., specificity level of motivation) moderate the strength of this mediation effect. The literature search provided 55 studies reporting 714 effect sizes. Significant mediation paths were found for all motivational constructs and had effect sizes ranging from $r_{MED} = .046$ (95% CI [.037, .056]) for extrinsic motivation to $r_{MED} = .199$ (95% CI [.027, .373]) for interest. Moderation analyses revealed that heterogeneity in effect sizes could be explained by the specificity level of motivation and the types of knowledge and interest. The results of this meta-analysis explain interindividual stability in knowledge and stress the role of motivation as an underlying mechanism. Low learner motivation can sometimes be improved by increasing prior knowledge, and low knowledge can sometimes be improved by increasing motivation.

6.2 Introduction

Prior domain-specific knowledge is assumed to be the strongest predictor of achievement and performance (Greve et al., 2019; Thompson & Zamboanga, 2003), yet identifying the processes responsible for this association remains an objective of current educational psychology research. In a systematic review of meta-analyses, Schneider and Preckel (2017) identified over 100 variables related to achievement in higher education, among which prior achievement was one of the strongest predictors. High-achieving students maintain their level, whereas low-achieving students do not seem to catch up with their more successful peers. What helps the former students, and why do the latter seem not to overcome their difficulties? These questions address the interindividual stability of knowledge within groups of learners over time and the underlying mechanisms. While some cognitive explanations have been uncovered for this phenomenon, such as the increased processing of information through chunking (Gobet, 2005) and optimized attention allocation (Gegenfurtner et al., 2011), there has been scattered research on non-cognitive explanations so far. Motivational research, which plays a vital role in educational psychology, could contribute to answering this question based on its large body of research (Murphy & Alexander, 2000). However, it provides only isolated explanations for the relation between prior knowledge and learning achievement, which remain to be synthesized.

Uncovering mechanisms of prior knowledge on learning outcomes is pivotal: The promotion of knowledge is the essential target of education and leads to several desirable outcomes associated with a fulfilling life, such as occupational success (Judge et al., 1995), well-being (Bücker et al., 2018), and health outcomes (Hahn & Truman, 2015; Simonsmeier et al., 2022). In instructional contexts, referring to students' prior knowledge is an effective method (Amadieu et al., 2015), but it could be used more purposefully if its mechanisms were better understood.

6.2.1 Knowledge

Researchers have tried to conceptualize knowledge and its utilization in different ways (Anderson & Krathwohl, 2001; Bloom, 1956; de Jong & Ferguson-Hessler, 1996). Perhaps the best-known attempt to describe knowledge comes from de Jong and Ferguson-Hessler (1996), with their distinction between types and qualities. They specify four types of knowledge that learners can use in problem-solving, namely situational, conceptual, procedural, and strategic knowledge. *Situational knowledge* contains knowledge about the relevant features of a situation, which can include aspects of conceptual or procedural knowledge. *Conceptual knowledge* refers to knowledge about facts, concepts, and principles that apply in a particular domain and provide information for solving a problem. The actions or manipulations required for the solution characterize *procedural knowledge*. Knowledge about the organization of phases in the problem-solving process, i.e., a general plan of action, is referred to as *strategic knowledge*.

Besides the types of knowledge, de Jong and Ferguson-Hessler (1996) suggest five different qualities of knowledge in use. Some directly particularize the knowledge types, others define the interrelationships between knowledge types, and sometimes the boundaries may overlap. First, *knowledge level* refers to the degree to which a knowledge type is thoroughly organized, elaborated, and well-understood in problem-solving. Superficial knowledge is associated with pure memorization and is less amenable to critical judgment and abstraction than deep knowledge. Second, the *structure of knowledge* describes whether the knowledge components within a domain are well or poorly connected. Structured knowledge can be achieved through efficient chunking and is observable in the knowledge of experts (Gobet, 2005). Third, the *use of knowledge* in problem-solving can also be more or less automated. Whereas novices perform their actions slowly and consciously, experts rely on automatic routines that allow them to interpret situations as a whole. Fourth, the *modality of knowledge* regards whether knowledge is stored in memory as a series of abstract sentences or images containing perceptual details. Depending on the situation, one may be more appropriate than the other. Last, the *generality of knowledge* describes the scope of the content. Domain-specific knowledge applies to specific situations and can hardly be transferred to other situations. Domain-general knowledge includes heuristics or strategies that learners can apply to multiple situations.

In a recent approach to conceptualize prior knowledge in the reading domain, McCarthy and McNamara (2021) proposed the *Multidimensional Knowledge in Text Comprehension framework*, which can easily be adapted for prior knowledge research in general. Their framework comprises four dimensions: amount, accuracy, specificity, and coherence, with each dimension affecting how information during learning is processed. The *amount* of knowledge refers to the quantity of information relevant to the subject matter. Learners with more prior knowledge are better at lower-level word processing (Priebe et al., 2012), making inferences (Kintsch, 1998), activating relevant information, and ignoring irrelevant information (McNamara & McDaniel, 2004). They also tend to use more sophisticated strategies than subjects with less prior knowledge (Cromley & Azevedo, 2007). The dimension of *accuracy* describes whether the knowledge is more or less accurate. While it is easy to identify some pieces of knowledge as misconceptions or as truly correct, other pieces of knowledge are harder to categorize, for example, when the correctness of information depends on its context. Incorrect prior knowledge poses a threat to correct knowledge acquisition because it is resistant to change (Smith III. et al., 1993), leads to the overconfidence of subjects in their responses in knowledge tests (Li et al., 2004), and remains a misconception when previously or later acquired correct knowledge cannot be used for knowledge integration (Clark & Linn, 2013). The *specificity* dimension refers to the subject matter of knowledge and is related to the generality knowledge quality by de Jong and Ferguson-Hessler (1996). The authors suggest a taxonomy in which the domain-specificity of the content is organized into seven levels, from general knowledge to knowledge of certain topics (McCarthy & McNamara, 2021). Knowledge in the form of general knowledge refers to academic or personal domains and is the most comprehensive. Instruments such as the Woodcock-Johnson Academic Knowledge test (Wendling et al., 2009) capture the knowledge of academic domains such as science, social studies, and humanities. These domains (e.g., science) can be subdivided into sub-domains (e.g., life science and physical science), which in turn include subjects (e.g., physics), sub-subjects (e.g., mechanics), and, at the most specific level, certain topics (e.g., momentum). The fourth dimension *coherence* describes the relations between the knowledge elements, which may be more interconnected and embedded in an overarching structure or somewhat fragmented and isolated.

For the present study, we define knowledge as information stored in memory, which includes declarative knowledge about facts and concepts as well as procedures. We consider knowledge at the specificity level of domain knowledge and below (McCarthy & McNamara, 2021) and encompass both correct and incorrect knowledge as our object of investigation. We define prior knowledge as the knowledge present in a person's long-term memory at the onset of learning (Dochy & Alexander, 1995). We mainly focus on the amount of prior and later knowledge instead of the quality. We do not consider the coherence of knowledge elements for this meta-analysis, as it would be necessary to analyze the wording of the knowledge instrumentation and the instructional material used in the primary studies, which are often not accessible in the articles.

6.2.2 The Impact of Prior Knowledge on Learning Outcomes

Simonsmeier et al. (2021) empirically tested the predictive power of prior knowledge on learning achievement in a meta-analysis of 493 studies. In detail, they conducted a literature search to capture study designs with knowledge assessments at two measurement points. They retrieved the correlations between the knowledge pre-test and the knowledge post-test scores to compute a pooled effect size and found a high correlation between pre-test and post-test knowledge, indicating that prior knowledge is a powerful predictor for subsequent knowledge and learning achievement. This could additionally be demonstrated in randomized controlled studies suggesting that the underlying relation is causal. Controlling for the influence of intelligence did not lead to significant decreases in the effect size. However, the large heterogeneity of effect sizes suggested other influences as well.

It is important to stress that the relation between pre-test and post-test knowledge does not reflect the relation between prior knowledge and learning of new material. Instead, it tells how the interindividual differences among groups of learners remain stable over time. Consider a sample of 100 participants taking a knowledge test on two occasions. Some of them do very well on the knowledge test, others make some mistakes, and some can barely answer a question correctly. This results in two rank orders of test scores, one for the pre-test and one for the post-test. Correlating the participants' pre-test and post-test scores shows how the test scores' rank orders remain stable. In a case where the correlation coefficient is 1, the rank order in the post-test scores is the same as in the pre-test. If the correlation coefficient were -1, the ranking in the post-test would be exactly the opposite of the ranking in the pre-test, with the person with the lowest test score in the prior knowledge test having the highest test score in the subsequent test and the best-performing person the prior knowledge test performing worst in the post-test. Note that the correlation coefficient also does not indicate whether the participants' performance improved or worsened over time. Thus, positive correlations between prior knowledge and post-test knowledge reflect the stability of individual differences in learners' knowledge but not how prior knowledge affects learning or how much was learned.

Simonsmeier et al. (2021) suggested that prior knowledge affects learning processes, which in turn influence learning outcomes. Therefore, these learning processes act as mediators between prior knowledge and learning outcomes. We hypothesize that motivational constructs may also function as mediators because they can both be predicted by prior knowledge and be predictors of learning outcomes, as we will describe in the following section.

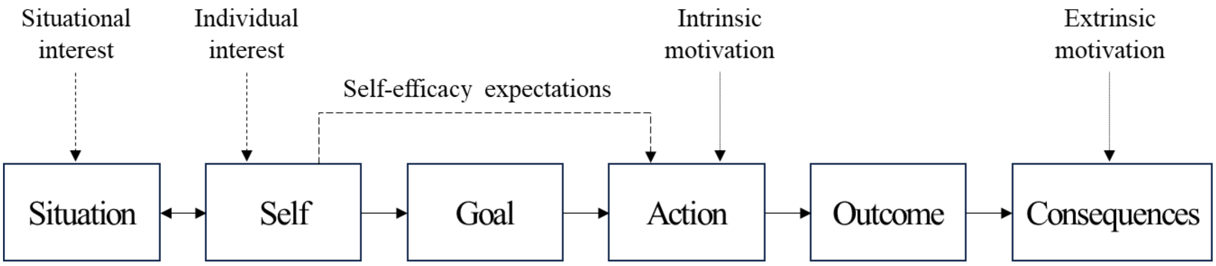
6.2.3 Key Constructs of Motivation

Vu et al. (2022) provided a framework for how achievement affects motivation and vice versa. The authors suggested the possibility that motivation influences achievement via two routes: The first route is through the number of academic behaviors, such as increased effort and persistence. The second route includes increased quality of academic behaviors, such as effective learning and metacognitive strategies. Similarly, achievement may affect motivation via two routes. Past achievement leads to perceptions of self-efficacy and control, thus affecting future expectations and the intrinsic value assigned to learning activities. Then, if a high value is attached to academic behaviors, it can create a rewarding experience of flow. Note that Vu et al. (2022) focused on prior achievement, including objective knowledge tests and grades and subjective assessments of teachers or learners. Such subjective assessments tend to inflate the effects associated with motivational constructs (Howard et al., 2021).

Motivation is an umbrella term for different motivational constructs with unique research traditions. Urhahne and Wijnia (2023) integrated theories of motivation in education in an integrative framework, including expectancy-value theory, social cognitive theory, self-determination theory, interest theory, achievement goal theory, and attribution theory. This model contains "the determinants and course of motivated action" (Urhahne & Wijnia, 2023, p.45) and grounds on the general model of motivation (Heckhausen & Heckhausen, 2018). Figure 26 outlines the fundamental aspects of the model with the main motivational constructs of this meta-analysis sorted into it. We focus on a selection of motivational constructs with the largest bodies of research, namely interest, self-concept, self-efficacy, intrinsic motivation, and extrinsic motivation. In the following section, we provide short descriptions of these constructs and outline how they both can affect and be affected by knowledge.

Figure 26

Integrative Framework of Theories of Motivation in Education



Note. Integrative framework of theories of motivation with relevant constructs sorted into the basic model based on interest theory, social cognitive theory, and self-determination theory (own representation, adapted from Urhahne & Wijnia, 2023, pp. 3, 10, 12, 15).

6.2.3.1 Interest. Interest has several characteristics: It is content or object-specific, involves a person-environment-interaction, has both cognitive and affective components, may be temporarily unconscious, and has a physiological and neurological basis (Renninger & Hidi, 2011). Models of interest development distinguish between situational and individual interest (Hidi & Renninger, 2006; Krapp, 2007). Situational interest is triggered by environmental characteristics that catch an individual's attention. Over time, if situational interest is maintained, it can transform into individual interest. This motivational disposition leads to self-initiated engagement with the task. Although being two separate constructs, knowledge and interest are highly related (Tobias, 1994). In the model of domain learning, Alexander et al. (1995) outlined the interplay between forms of interest and knowledge in the development of expertise. During the early stages of expertise development, when individual interest and knowledge in the domain are relatively low, situational interest is mainly responsible for the engagement with a certain task. As learners acquire more knowledge in the domain, their individual interest grows, and the influence of situational interest on engagement weakens.

Evidence suggests that knowledge and interest are related bidirectionally (Rotgans & Schmidt, 2017). Knowledge as a cause of interest, however, is rarely investigated. For example, in a longitudinal study, Zhang et al. (2016) used growth curve analyses to find that prior knowledge affected the development of interest in learning declarative and procedural knowledge about physical activity. Interest as a predictor of learning outcomes has been studied more frequently. In a meta-analysis, Schiefele et al. (1992) found that interest predicted achievement well in multiple subject areas.

6.2.3.2 Self-concept. In general, self-concept refers to an individual's perception of themselves as shaped by experiences with the environment (Shavelson et al., 1976). In academic contexts, self-concept refers to a "mental representation of one's academic abilities in general and in different academic domains" (Arens et al., 2021, p. 2). In addition to the cognitive self-evaluation, it has an affective component resulting from comparing one's attributes and competence to standard (Arens & Hasselhorn, 2015; Bong & Clark, 1999). Typical measures of academic self-concept refer to how well students do in a particular subject or how quickly they learn the material (e.g., Marsh et al., 2014). Although different structures of academic self-concept exist, most researchers agree on its multidimensional and hierarchical structure (Arens et al., 2021).

Due to the nature of academic self-concept being a self-evaluation of achievement, it is closely related to knowledge. Both grades and objective achievement tests affect later measures of self-concept, with grades having a slightly greater impact (Helmke & Van Aken, 1995). In some cases, academic self-concept is as powerful as intelligence in predicting test performance (Lauermann et al., 2020). In turn, self-concept predicts later academic achievement, especially when the assessments of self-concept are formulated more specifically: There are stronger effects for task-specific self-concept than for academic and general self-concept (Choi, 2005), especially when they refer to the same subject area (Valentine et al., 2004). A meta-analysis of the longitudinal relation between academic self-concept and achievement found evidence for small reciprocal effects (Wu et al., 2021).

6.2.3.3 Self-efficacy. Self-efficacy refers to an individual's belief in his or her capability to perform certain actions (Bandura, 1977). Like self-concept, it can be referenced to contexts with varying degrees of specificity (general, academic, or task-specific; Choi, 2005). Academic self-efficacy refers to judgments of the capability to accomplish certain academic tasks at a certain level (Schunk, 1991). Academic self-efficacy and academic self-concept share many similarities. Some researchers investigated whether they represent different constructs (Bong & Skaalvik, 2003): Both are based on a self-assessment of one's competence; they are domain-specific and multi-dimensional. Additionally, they have similar predictive power for different outcomes. However, self-efficacy does not include an affective component and pertains to the achievement of specified future goals, thus being less stable over time than self-concept.

From a theoretical perspective, it seems plausible that prior knowledge and achievement in a domain are sources of academic self-efficacy. However, few studies examined the impact of prior knowledge on self-efficacy (Honicke & Broadbent, 2016). However, it appears safe to assume that past performance in academic contexts affects self-efficacy (Brown et al., 2008). On the other side, the effects of self-efficacy on learning and achievement are better documented. Self-efficacy for learning and achievement predicts grade point average and course grades (Crede & Philips, 2011) and shows the second-largest impact on achievement in higher education in a review of meta-analyses (Schneider & Preckel, 2017). Possible explanations for this relation are the increased use of sophisticated monitoring strategies, effort, and adequate goal setting with high levels of self-efficacy (Cheng & Chiou, 2010; Kassab et al., 2015; Moos & Azevedo, 2009). The mediation hypothesis for the prior knowledge-knowledge relationship holds true, at least for high school and college GPAs (Brown et al., 2008).

6.2.3.4 Intrinsic Motivation. Intrinsic motivation is defined as the desire to perform a certain action because the action itself is perceived as interesting, challenging, or otherwise enjoyable (Ryan & Deci, 2000b). According to self-determination theory, intrinsic motivation is based on three basic needs of experiencing (competence, autonomy, and relatedness) that determine motivation and behavior (Deci & Ryan, 2000). Over time, actions that were originally initiated because of an external reward or punishment (e.g., receiving an ice cream for completing homework or being scolded for not doing so) may become internalized so that the performance of the action is consistent with one's mastery goals and personal values that are identified as part of one's self (e.g., doing homework because it is enjoyable; Ryan et al., 2021).

It seems plausible to assume that prior knowledge can be a source of experiencing competence and therefore affects intrinsic motivation. However, only one record investigated this relation (Achmetli et al., 2019). In this study, both conceptual and procedural knowledge led to experiences of competence in a mathematical multiple-solution task. Experiences of competence then predicted conceptual and procedural knowledge at the post-test. The effects of intrinsic motivation on academic performance have been investigated more thoroughly. A meta-analysis by Cerasoli et al. (2014) found a moderate correlation between intrinsic motivation and performance. The proportion of variance explained was larger when the quality of performance (e.g., creativity) rather than quantity (e.g., points in a test) was considered.

6.2.3.5 Extrinsic Motivation. Extrinsic motivation is characterized by the desire to perform a certain action to receive an expected reward or to prevent a negative consequence, making it a rather instrumental behavior (Ryan et al., 2021). Several forms of extrinsic motivation exist with different degrees of internalization ranging from complete external regulation to the involvement of personal goals. They can co-exist within learners (Ratelle et al., 2007), leading to different academic outcomes. Extrinsic motivation is related to performance goals in goal-setting frameworks and to the surface processing of information

(Elliot et al., 1999).

Prior knowledge may affect the forming of extrinsic motivation via increased expectations of success elicited by previous experiences of success and achievement (Eccles & Wigfield, 2020). This relation, however, was rarely addressed. Liu and Hou (2017) found that prior math achievement predicted test-taking motivation, the motivation to perform well on a given test in mathematics. Again, the impact on learning outcomes was investigated more often. The meta-analysis by Cerasoli et al. (2014) found that extrinsically motivated behavior was related to performance with higher proportions of variance explained for quantitative outcomes than qualitative outcomes.

6.2.4 Influences of Potential Moderators

Due to the broad field of motivational research, we expect high heterogeneity among the effect sizes reflecting associations between prior knowledge, motivation, and learning outcomes. Heterogeneity can emerge from features of the primary studies, such as different instrumentations, knowledge domains, or features of learning phases. Moderator analyses identify such influences of third variables and explain to which degree the results are affected by these. They are based on theoretical assumptions as well as on exploratory examinations.

6.2.5 Aims of the Current Study

Although motivational research has significantly contributed to predicting learning outcomes, it has not yet been identified as a process variable mediating the relation between prior knowledge and learning outcomes using meta-analytic methods. Research syntheses such as meta-analyses can provide insights into research questions, although they have never been explicitly addressed in primary studies. In particular, the effects of prior knowledge on later assessments of motivation were rarely the focus of primary studies. However, they are frequently reported as side results and can be used to determine an overall effect size across studies. Combining the effects of prior knowledge on motivation and the effects of motivation on learning outcomes allows the investigation of mediation effects.

In addition, if these mediation effects can be demonstrated, we test for moderating influences of third variables. As described in the motivation section (Chapter 6.2.3), we expect larger effect sizes on the task-specific level of motivation than for motivation referring to the general and academic levels (Choi, 2005). Furthermore, we conduct exploratory analyses on moderating influences of knowledge type (declarative vs. procedural) and interest type (situational vs. individual). We formulate three hypotheses:

1) Hypothesis 1: Prior knowledge is a predictor of future interest, self-concept, self-efficacy, intrinsic motivation, and extrinsic motivation.
2) Hypothesis 2: The motivational constructs considered in Hypothesis 1 act as mediator variables for the relation between prior knowledge and learning outcomes.
3) Hypothesis 3: The indirect effects are moderated by the specificity levels of motivation. Task-specific motivation shows larger effect sizes than motivation on the general or academic level.

6.3 Method

6.3.1 Literature Search

The literature search was conducted in the databases PsycINFO and ERIC in September 2020. For the search, the default settings of the databases were used, which, among other components, include title, abstract, and keywords of potentially relevant articles. 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*") and knowledge) or "prior knowledge" or "knowledge change" or "knowledge gain*") and ("motivation*" or "self-efficacy" or "self-concept" or "interest" or "confidence")). This search string was intended to not only capture studies that explicitly focused on the role of prior knowledge and motivation in learning as such. Rather, the search string should also capture designs that assessed knowledge at two measurement points and motivation at at least one measurement point. (e.g., pre-post educational interventions aiming at the acquisition of knowledge and skills, including motivation as a predictor).

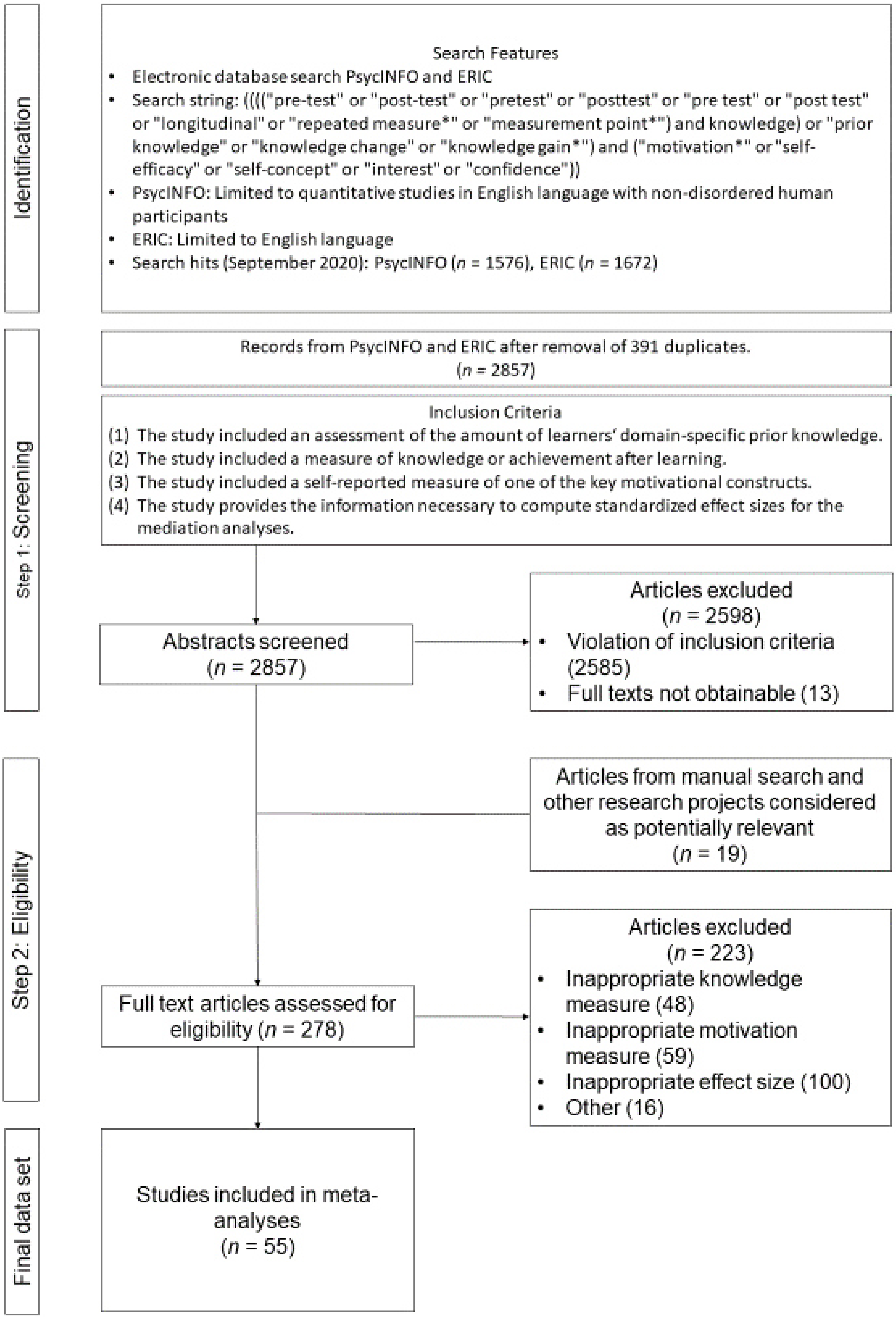
Search results were limited to studies in English in PsycINFO and ERIC. Additionally, in PsycINFO, the results were restricted to quantitative studies with human non-disordered participants. Unpublished documents were included to reduce the probability of publication bias (Dickersin, 1990). After the removal of duplicates, the search results comprised 2857 research articles. Additionally, 19 articles from manual searches and other research projects were considered potentially relevant. A summary of the literature search process and the inclusion criteria is depicted in Figure 27.

6.3.2 Inclusion Criteria

Based on the meta-analysis by Simonsmeier et al. (2021), studies that met the following four criteria were included in the meta-analysis: (1) The study included an assessment of the amount of learners' domain-specific prior knowledge. Only objective quantitative measures of domain-specific prior knowledge were included. Self-assessments, composite scores from more than one domain, and measures of crystallized intelligence, abilities, achievement, or meta-cognitive knowledge were excluded. (2) The study included a measure of knowledge or achievement after learning. We excluded learners' self-ratings of their learning outcomes and knowledge of a different domain than the prior knowledge domain. (3) The study included a self-reported measure of one of the key motivational constructs, namely interest, self-concept, self-efficacy, and intrinsic or extrinsic motivation. (4) The study provided the information necessary to compute standardized effect sizes for the mediation analyses. This required information on all three relations between prior knowledge, post-test knowledge, and a motivational construct.

Figure 27

Flow Chart for the Literature Search and Inclusion Process (Study 2)



The 2857 articles from the literature search were screened to determine if they met the inclusion criteria. After a coder training, the first and third authors screened 100 titles and abstracts for inclusion. The intercoder agreement for this subset of articles was 81%. Disagreements were resolved by discussion. The first author screened the remaining titles for inclusion in the next step. In sum, 278 articles were obtained for further examination. In the next step, the first and second authors checked whether the first 100 articles fitted the meta-analysis following the inclusion criteria. The agreement for this step was 91%. After that, the first author screened the remaining articles for inclusion. In sum, 223 articles were excluded resulting in 55 studies included in the meta-analysis.

6.3.3 Data Coding

Study information was coded using predefined coding rules for coder training and the final extraction of study information with a standardized coding instrument. Table 8 lists the study and effect size characteristics considered for data extraction. The first author and the second author independently coded a random selection of 50 effect-size triplets consisting of the correlations between prior knowledge, the motivational measure, and post-test knowledge, as well as their corresponding characteristics from the full texts. The inter-rater agreement for the moderator variables ranged from 64% to 100%. The mean agreement across all moderators was 89%. Disagreements were resolved through discussion. After that, the first author coded the remaining effect size triplets. If a study included multiple measurement points with unique intermediary learning phases, all possible relations between the constructs were coded (e.g., T1 knowledge with T2 knowledge and T2 motivation as well as T1 knowledge with T2 motivation and T3 knowledge). However, correlations with a measure of motivation outside the knowledge pre-test and post-test time window were not coded (e.g., T1 motivation with T2 knowledge and T3 knowledge).

Table 8

Description of the Included Information (Study 2)

Moderator	Description
Knowledge characteristics (for both measurement points)	
Broad content domain	The mediation might differ based on the broad content domain of the assessment. We coded five broad content domains: *STEM*, *language*, *social sciences*, *health sciences*, and *other*.
Content domain	To test whether the mediation differs between studies analyzing different content domains, we coded seven content domains: *Mathematics*, *L1*, *biology*, *physics*, *chemistry*, *social sciences*, and *other*.
Specific content	We coded the specific content.
General knowledge type	We distinguished between *declarative* and *procedural* knowledge as general knowledge types.
Specific knowledge type	Four specific knowledge types were coded: *Facts*, *concepts*, *cognitive skills*, and *mixed*.
Number of items	We coded the mean number of items.
Response format	Three levels were distinguished in the response format: *Multiple choice*, *open*, and *mixed*.
Reliability	We coded the reliability of the knowledge assessment.
Knowledge characteristics at measurement point two	
Achievement	Whether the knowledge assessment at measurement point two was measured via *achievement* or a *knowledge test* was coded.
Feedback	We coded *yes* when feedback was given and *no* when no feedback was provided.
Grading	To assess whether grading influenced the mediation, we coded *yes* when students received grades and *no* when they did not.
Same Test T1-T2	We coded whether the same test was given at both measure points with *yes* or *no*.
Mediator characteristics	
Motivational construct	We coded six levels of motivational constructs: *Confidence*, *self-efficacy*, *self-concept*, *intrinsic motivation*, *extrinsic motivation*, and *interest*.
Type of interest	We differentiated between *personal interest* and *situational interest*.
Specificity	We coded the specificity of the corresponding motivational construct as *general*, *academic*, and *specific*.
Mediator time	The measurement point of motivation could influence the mediation. Therefore, we coded six time points of assessment: *Before T1*, *At T1*, *Between T1 and T2*, *At T2*, *After T2*, and *no information*.
Reliability	We coded the reliability of the motivation assessment.

Table 8 (continued)

Sample characteristics	
Country	To test influences of the country, we coded the country of the primary study or, if not stated, of the first author.
Age group	We distinguished between different age groups. We coded the following levels: *Kindergarten / Preschool*, *Primary education*, *Secondary education*, *Higher education*, and *Other*.
N of the sample	We coded the total N of the sample.
Age	We coded the sample mean age of the learners (in years).
Methodological study characteristics	
Grey literature	We coded whether the primary study is grey literature with *yes* or *no*.
Study design	We coded whether the primary study was an *experiment*, *quasi-experimental*, or *longitudinal*.
Prior knowledge type	We distinguished between preexisting prior knowledge or manipulation with two levels: *Preexisting* and *Randomization*.
Motivation manipulation	We coded whether motivation was manipulated (*Yes*) or not (*No*).
Mean days between intervention and transfer test	We coded the mean number of days between the two measurement points.
Intervention	To test whether lessons or learning environments influence the mediation, we coded two levels: *yes* and *no*.
Mean number of sessions	We coded the mean number of sessions.
Mean duration of sessions	We coded the mean duration in minutes of sessions.

Note. Overview of included moderators and the coded categories. STEM - Science, technology, engineering, and mathematics; L1 – First language.

6.3.4 Requests for Missing Information in the Articles

Some of the included full texts formally met all inclusion criteria except for reporting the effect sizes necessary for the computations of the meta-analysis. For this reason, we sent 87 emails to the authors of those articles asking them to provide the missing correlations or the data so the effect sizes could be calculated. This procedure obtained 280 additional effect sizes from 13 studies.

6.3.5 Statistical Analyses

6.3.5.1 Meta-Analytic Integration. To determine the pooled effect sizes for the unique paths in the mediation model, we followed the product-of-coefficients strategy by Baron and

Kenny (1986). The path pointing from prior knowledge to motivation is called b_1, and the path pointing from motivation to learning outcomes is called b_2. The correlation between prior knowledge and post-test knowledge represents the *total effect* in the mediation model. The *indirect effect* in the mediation model reflecting the mediation was calculated by multiplying b_1 with b_2 for each study. Subtracting the indirect effect from the total effect yields the remaining *direct effect* of prior knowledge on post-test knowledge controlled for motivation. To investigate Hypothesis 1, only the b_1 paths were considered in which motivation was assessed between T1 and T2 or at T2. For Hypothesis 2, we examined the indirect effect of each motivational mediator. The effect sizes were inverted whenever low scores represented high motivation to reflect the actual relations adequately.

We applied a random-effects model using the *metafor* package (Viechtbauer, 2010) in R (R Core Team, 2022). Most studies reported more than one relevant effect size because multiple measurement points or samples provided valuable information. These effect sizes are statistically dependent. For this reason, we employed robust variance estimation (Hedges et al., 2010) to ensure that dependent effect sizes did not affect the validity of the meta-analytic results. In addition, we corrected for imperfect measurements of the knowledge and motivation instruments using the formula from Schmidt and Hunter (2015, p. 144) and the reliability coefficients stated in the articles. If the reliability of the measures was not stated in the articles, no correction was made.

6.3.5.2 Moderator Analyses. We calculated the mean effect sizes separately for each level within a moderator category and determined their confidence intervals and indicators. For moderator levels with less than four degrees of freedom, we reported only the mean estimates but not the confidence intervals because the test of significance is impaired by the small number of observations (Tanner-Smith et al., 2016). Each moderator level was dummy-coded to be entered as a predictor in the mixed-effect models. One moderator level served as a reference category in each analysis. It was compared with the other moderator levels to identify significant differences. In addition, we computed the amount of explained variance R^2 for each regression model.

6.3.5.3 Publication Bias. Results from meta-analyses can be compromised by publication bias. Large significant effect sizes tend to be published more than insignificant or small effect sizes. Therefore, we visually and statistically checked for publication bias using funnel plots and Egger regression (Egger et al., 1997) with the *metafor* package (Viechtbauer, 2010).

6.4 Results

6.4.1 Characteristics of Included Studies and Effect Sizes

The 55 included studies provided 714 effect size triplets (for the list of included studies, see Appendix F). All effect sizes were bivariate correlations. The correlations between prior knowledge and learning outcomes ranged from r = -.09 to r = .92, whereas the correlations between prior knowledge and motivation ranged from r = -.22 to r = .90, and the correlations between motivation and learning outcomes ranged from r = -.16 to r = .94. The primary studies have been published between 1994 and 2019 with a median of 2014. Seven of the 55 studies were grey literature (13%). The dataset comprised 56382 participants from 80 different samples and 14 different countries. The mean age of the samples ranged from 4.35 years to 44.2 years, with a median of 14.7, indicating that most of the participants were school students.

The sample sizes ranged from 22 to 16110 participants, suggesting that small local and large national studies were included in the dataset. The time between the measurement of prior knowledge and learning outcomes ranged from seven to 2920 days, with a median of 180 days. Subdivided by motivational constructs, there were 31 articles for interest, eight for self-concept, 21 for self-efficacy, 15 for intrinsic motivation, and six for extrinsic motivation. The number of effect size triplets was 310 for interest, 97 for self-concept, 208 for self-efficacy, 74 for intrinsic motivation, and 25 for extrinsic motivation.

6.4.2 Main Meta-Analytic Results

6.4.2.1 Hypothesis 1: Effects of Prior Knowledge on Motivation. Prior knowledge significantly predicted motivation at a later measurement point across all the motivational constructs considered (r_{b1}). Prior knowledge most strongly predicted interest with r_{b1} = .409 (95% CI [.382; .436]) followed by self-concept with r_{b1} = .306 (95% CI [.268; .344]). The correlation between prior knowledge and extrinsic motivation was smaller (r_{b1} = .176, 95% CI [.117; .234]), similar to intrinsic motivation (r_{b1} = .168, 95% CI [.110; .225]) and self-efficacy (r_{b1} = .139, 95% CI [.105; .172]). For all the pooled correlations, the heterogeneity Index I^2 was high (> 90%, Higgins et al., 2003), indicating that a large proportion of variance in the effect sizes was due to influences of moderator variables (Borenstein et al., 2017).

6.4.2.2 Hypothesis 2: Analyses of Mediation Effects. As shown in Table 9, all paths considered in the mediation model were significant, including the indirect mediation paths. The estimates of the indirect path coefficients (r_{MED}) were small but significant. The largest estimate was for interest as a mediator with r_{MED} = .199 (95% CI [.027; .373]), followed by self-concept with r_{MED} = .142 (95% CI [.051; .234]). The indirect effects of self-efficacy with r_{MED} = .077 (95% CI [.063; .091]), intrinsic motivation with r_{MED} = .050 (95% CI [.024; .076]), and extrinsic motivation with r_{MED} = .046 (95% CI [.037; .056]) were slightly smaller.

The pooled correlation between prior knowledge and learning outcomes, the total path (r_{TOT}), across all studies was r_{TOT} = .53 (95% CI [.514; .555]), which is following the results from Simonsmeier et al. (2021) concerning the stability of individual differences. Heterogeneity among the effect sizes was high, with I^2 = 99.96%. Depending on the subset of effect sizes for each motivational construct, the total effect ranged from r_{TOT} = .287

Table 9

Model Results and Heterogeneity Indices for Each Path Considered in the Mediation Models Separated by Mediators

				Model Results							Heterogeneity Indices							
Mediator	Mediation path	*j*	*k*	*r*	*SE*	*t*	*p*	CI lb	CI ub	sig	τ^2	SE	τ	I^2	H^2	*Q*	df	*p*
Interest	*b1*	310	31	.351	0.10	3.45	< .001	.143	.559	**	.045	0.00	.213	95.07	20.26	6432.92	309	< .001
	b2	310	31	.421	0.12	3.4	< .001	.168	.674	**	.064	0.01	.253	97.12	34.78	10480.35	309	< .001
	Total	310	31	.617	0.07	8.48	< .001	.468	.765	***	.045	0.00	.211	99.73	370.90	35991.46	309	< .001
	Indirect	310	31	.199	0.09	2.33	.03	.024	.373	*	.024	0.00	.156	89.01	9.10	2828.52	309	< .001
	Direct	310	31	.424	0.05	8.87	< .001	.326	.521	***	.029	0.00	.169	95.75	23.55	14103.66	309	< .001
Self-Concept	*b1*	97	8	.322	0.05	6.77	< .001	.209	.434	***	.024	0.00	.156	98.57	69.93	5849.30	96	< .001
	b2	97	8	.375	0.05	8.23	< .001	.267	.482	***	.028	0.00	.167	99.5	200.42	82701.68	96	< .001
	Total	97	8	.727	0.06	12.85	< .001	.593	.861	***	.035	0.01	.188	99.97	3846.86	61976.95	96	< .001
	Indirect	97	8	.142	0.04	3.68	.01	.051	.234	**	.018	0.00	.133	97.55	40.85	2580.23	96	< .001
	Direct	97	8	.585	0.04	13.99	< .001	.486	.684	***	.034	0.01	.185	99.65	287.33	25554.83	96	< .001
Self-efficacy	*b1*	208	21	.213	0.02	12.57	< .001	.178	.248	***	.053	0.01	.231	99.93	1461.45	53073.27	207	< .001
	b2	208	21	.285	0.03	9.23	< .001	.220	.349	***	.101	0.01	.317	99.99	17439.33	50279.97	207	< .001
	Total	208	21	.287	0.07	4.08	< .001	.140	.433	**	.043	0.01	.208	99.86	696.75	35094.69	207	< .001
	Indirect	208	21	.077	0.01	11.56	< .001	.063	.091	***	.01	0.00	.100	79.28	4.83	875.68	207	< .001
	Direct	208	21	.213	0.07	2.91	.01	.060	.365	**	.044	0.01	.210	99.68	311.89	45532.96	207	< .001

Table 9 (continued)

Mediator	Mediation path	*j*	*k*	Model Results							Heterogeneity Indices							
				r	*SE*	*t*	*p*	CI lb	CI ub	sig	τ^2	SE	τ	I^2	H^2	*Q*	df	*p*
Intrinsic Motivation	*b1*	74	15	.180	0.03	6.23	< .001	.118	.242	***	.023	0.01	.152	96.59	29.3	1638.64	73	< .001
	b2	74	15	.218	0.02	10.18	< .001	.172	.264	***	.012	0.00	.109	93.32	14.96	1568.49	73	< .001
	Total	74	15	.582	0.10	5.67	< .001	.362	.802	***	.079	0.01	.282	99.96	2516.17	12396.77	73	< .001
	Indirect	74	15	.050	0.01	4.17	< .001	.024	.076	**	.002	0.00	.046	68.03	3.13	433.43	73	< .001
	Direct	74	15	.539	0.10	5.3	< .001	.321	.757	***	.075	0.01	.273	99.92	1176.52	14795.68	73	< .001
Extrinsic Motivation	*b1*	25	6	.165	0.05	3.04	.03	.025	.305	**	.012	0.00	.111	95.7	23.27	222.5	24	< .001
	b2	25	6	.158	0.05	2.99	.03	.022	.294	**	.012	0.00	.112	95.36	21.57	190.55	24	< .001
	Total	25	6	.679	0.11	6.17	< .001	.396	.962	**	.039	0.01	.197	99.8	512.5	904.73	24	< .001
	Indirect	25	6	.046	0	12.46	< .001	.037	.056	***	0	0	0	0	1	6.52	24	1
	Direct	25	6	.649	0.10	6.76	< .001	.402	.896	**	.030	0.01	.174	99.62	260.64	680.21	24	< .001

Note. Overview of model results and heterogeneity indices; *b1* - *b1*-path (prior knowledge to motivation), *b2* - *b2*-path (motivation to learning outcomes), Total - Total path (prior knowledge to learning outcomes), Indirect - Mediation path (b1 * b2), Direct - Direct path (total path minus indirect path), *j* - Number of effect sizes, *k* - Number of studies, *r* - Mean correlation coefficient, *SE* - Standard error, *t* - *t*-value, *p* - p-value, CI lb - Confidence interval lower bound, CI ub - Confidence interval upper bound, sig - Significance (* = $p < .05$, ** = $p < .01$, *** $p < .001$), τ^2 - tau squared, τ - tau, I^2 - *I* squared, H^2 - *H* squared, *Q* - *Q*-statistics, df - degrees of freedom.

(self-efficacy) to $r_{TOT} = .727$ (self-concept). All direct paths were significant, implying that none of the motivational constructs fully accounted for the association between prior knowledge and learning outcomes.

6.4.3 Moderator Analyses

6.4.3.1 Hypothesis 3: Specificity Level of Motivation. A moderator analysis of the indirect effects averaged over all motivational constructs revealed no significant differences between the academic-level and the task-level of specificity ($p = .23$). There was only one case of general specificity and in another case, the level of specificity could not be determined from the article. However, considering individual motivational constructs, some mediator levels resulted in significant differences: When the instrument measuring interest was formulated on the academic level, the estimate was $r_{MED} = .025$ (95% CI [-.025; .071]), whereas the estimate of the indirect effect for task-specific formulations of interest was $r_{MED} = .229$ (95% CI [.210; .248]). This difference was significant ($p < .001$). The indirect effect for self-concept was also significantly stronger when the instrument referred to (task-)specific contexts ($r_{MED} = .281$, 95% CI [.083; .219] for specific; $r_{MED} = .129$, 95% CI [-.112; .081] for academic, $p < .001$). However, this finding must be interpreted cautiously as fewer than four studies had reported specific self-concept. No significant differences between the levels of self-efficacy were found despite having sufficient different articles for a moderator analysis ($p = .513$). Not enough articles were available for the other motivational constructs to compute reliable results.

6.4.3.2 Exploratory Moderator Analyses. Averaged over the indirect effects of all motivational constructs, the estimate was significantly larger for declarative knowledge ($r_{MED} = .140$, 95% CI [.128; .153]) than for procedural knowledge ($r_{MED} = .112$, 95% CI [.087, .137], $p < .001$). Taking the subsets of mediators into account, significant differences also emerged for interest ($r_{MED} = .210$, 95% CI [.194; .233] for declarative knowledge; $r_{MED} = .030$, 95% CI [-.040; .100] for procedural knowledge, $p < .05$) and for self-concept ($r_{MED} = .086$, 95% CI

[.050; .127] for declarative knowledge; r_{MED} = .180, 95% CI [.147; .213] for procedural knowledge, $p < .001$). However, only two studies reported self-concept in connection with declarative knowledge, so these results cannot be fully trusted. There were no statistically significant differences between declarative and procedural knowledge in the indirect effects of self-efficacy. The same finding applies to intrinsic and extrinsic motivation, although there were not enough studies to fully confirm these relations.

The comparison of interest types revealed a significant difference between the indirect effects of individual interest and situational interest. On average, the indirect effect for individual interest was significantly stronger (r_{MED} = .208, 95% CI [.188; .227]) than for situational interest (r_{MED} = .107, 95% CI [.043; .171], $p < .01$). Twenty-seven studies reported effect size triplets for individual interest whereas eight studies reported situational interest, allowing this finding to be considered trustworthy.

6.4.4 Publication Bias

The funnel plots on each of the paths are depicted in Appendix G. The examination of the funnel plots and the Egger regression test pointed towards an asymmetry of the overall correlation between prior knowledge and learning outcomes, thus implying that large significant effect sizes were more likely to be published than small or not significant effect sizes. For interest, although the Egger test for asymmetry was significant for the b_1-path and the b_2-path, the visual inspection of the funnel plots does not imply that the results were biased towards larger effect sizes. Both the visual impression and the Egger tests for intrinsic motivation did not indicate a publication bias. However, the effect sizes for self-concept, self-efficacy, and extrinsic motivation might be affected by publication bias, as supported by significant results of Egger regressions.

6.5 Discussion

6.5.1 Main Findings

The aims of the current meta-analysis were (1) to investigate the predictive power of prior knowledge on selected motivational constructs, (2) to test the mediating influence of motivation on the relation between prior knowledge and learning outcomes, and (3) to identify moderating influences on this mediation. As a first important finding, prior knowledge significantly predicted all motivational constructs considered, namely interest, self-concept, self-efficacy, intrinsic motivation, and extrinsic motivation (r_{b1} = .139 for self-efficacy to r_{b1} = .409 for interest). All relations were positive, indicating that prior knowledge poses a resource for learners' motivation. This finding aligns with the framework proposed by Vu et al. (2022). The average effect sizes, however, varied depending on the motivational construct. Whereas the highest relations were found for interest and self-concept, the predictive quality of prior knowledge for later intrinsic and extrinsic motivation and self-efficacy was lower. To our knowledge, the current meta-analysis is the first to address the impact of objective prior knowledge as assessed with knowledge tests on later motivation. Previous studies focused on prior achievement, which is related, but not equal to prior knowledge because achievement also includes course grades and teachers' subjective assessments or learners' self-evaluations.

A second important finding of this meta-analysis was that motivation partly mediated the relation between prior knowledge and learning outcomes. Thus, successful students that scored high on a previous knowledge test before the actual learning phase performed better than their peers, partly because they were more motivated by the knowledge they already possessed. The strength of the mediating influence depended on the type of motivation being considered. Interindividual stability in knowledge test scores could be explained best by interest (r_{MED} = .199, 95% CI [.027; .373]) and self-concept (r_{MED} = .142, 95% CI [.051; .234]). The significant mediations of interest and self-concept are in line with two prior studies in which domain-specific curiosity, a construct closely related to interest, and mathematical self-concept mediated the relation between prior knowledge and achievement as well as learning outcomes (Watts et al., 2015; Witherby & Carpenter, 2021). The large confidence intervals for the indirect effects of interest and self-efficacy imply that the mediating influences could be considerably stronger or weaker. Nonetheless, self-efficacy, intrinsic motivation, and extrinsic motivation can also explain this issue. Due to the significant remaining direct paths controlled for the impact of motivation, the inspected motivational variables alone do not account entirely for the shared variance between prior knowledge and learning outcomes.

The insights from the moderator analyses show that heterogeneity among the mediation paths is partly due to methodological differences and construct-related features. Considering all motivational constructs, there was no moderating effect of the specificity level of motivation on the mediation path. However, an examination of the individual motivational constructs revealed that high specificity accompanies larger effect sizes for interest and probably for self-concept, too. For self-efficacy, no impact of specificity on the strength of mediation was found. Considering the findings from Choi (2005) that pointed towards increasing effect sizes with increasing levels of specificity for self-efficacy, the impact of specificity does not seem to be strong enough to affect the mediating role of self-efficacy. However, the low number of studies involved in the moderator analysis might have also affected this outcome. Additionally, the type of knowledge (declarative vs. procedural) affected the mediating role of motivation, albeit minimally, when considering all motivational constructs at once. The difference became considerably larger when only interest was examined. In addition, interindividual differences between learners could be better explained by individual interest than situational interest.

6.5.2 Limitations

We caution against making causal statements about the influence of prior knowledge on motivation and learning using the results of our meta-analysis. First, all effect sizes considered in our dataset reflected longitudinal relationships between the variables. The mediation paths in our results were significant, meaning there is a significant amount of shared variance between motivation and the two knowledge assessments at two measurement points, and multiplying the correlation coefficients b_1 and b_2 still leads to significant results. However, the question of whether the underlying explanation is causal or if motivation acts as a third variable explaining variance in both prior knowledge and learning outcomes cannot be answered with our methodology. We additionally coded effect sizes in which motivation was assessed with prior knowledge or learning outcomes at the same measurement point. From a theoretical perspective and with the support of the current results, we conclude that causal effects are plausible, but experimental designs would be needed for definite conclusions.

Second, our results suggest an explanation for why the rank orders of learners' test scores in a sequence of knowledge tests remain stable but do not provide evidence for the learning new information. In other words, the results from our study explain why successful students in prior knowledge tests are also more successful than their peers in a knowledge test after learning, but they do not explain that students with high prior knowledge learn more than their less knowledgeable peers because they are more motivated to do so. To make statements about learning gains, the analysis of correlations with learning gains would be necessary, but for mediation analyses on a meta-analytical level, there are still too few effect sizes. Note that neither correlations between prior knowledge and learning outcomes nor correlations between prior knowledge and learning gains allow for statements about how much knowledge learners acquire.

6.5.3 Theoretical Implications

Different forms of motivation partly accounted for the relation between prior knowledge and learning outcomes. Our meta-analysis supports the motivation-achievement cycle, as Vu et al. (2022) suggested. Motivational constructs related to both expectancy appraisals (such as self-efficacy) and value appraisals (such as interest) show significant mediating effects. Framed differently, the results suggest a rich-get-richer effect because high-achieving students can rely on their motivation formed by prior successes, whereas low-achieving students have to find other sources of motivation. Moreover, it seems complicated to close gaps between learners with much and little prior knowledge by promoting interest. Since situational interest is a mediator of comparatively little importance, sufficient prior knowledge would be necessary to develop a strong individual interest. Researchers interested in the magnitude of effects of prior knowledge on motivational constructs should therefore be careful when considering high or low degrees of prior knowledge in the sample.

Eccles and Wigfield (2020) demanded more research on the development of values assigned to tasks and the expectancy of success when engaging in these tasks. It is not yet well understood from which sources individuals draw their information to form task values and expectancies. Although our results do not provide causal explanations, they represent an essential first step to considering objective prior knowledge as one of such possible sources.

6.5.4 Practical Implications

Our results imply that the most successful students in a group of learners continue to be the most successful due to increased experiences of interest, self-concept, self-efficacy, as well as intrinsic and extrinsic motivation. Influential motivational theories such as the situated expectancy-value theory (Eccles & Wigfield, 2020) and self-determination theory (Ryan & Deci, 2002) suggest that one's experience of competence is essential in the emergence of motivation. Accordingly, low-performing students may not have acquired sufficient experience to develop these forms of motivation. In fact, low-achievers show different motivational patterns than high-achievers (McCoach & Siegle, 2001). For this reason, interventions supporting students' sense of competence to raise their expectations may help them catch up in their performance. One possible method in such interventions could be a form of feedback that emphasizes what knowledge and skills the learner already possesses. Even if the feedback is negative, it is possible to prevent lowering the learners' motivation (Fong et al., 2019). In a recent meta-analysis, Wisniewski et al. (2020) found that high-information feedback is more effective than simple forms of reinforcement and punishment. However, they also found that feedback has lower effects on motivational outcomes than on cognitive outcomes. If the information on learners' current state of knowledge was included, this would perhaps lead to increased effects on motivational outcomes of feedback.

6.5.5 Implications for Future Research

With the results of our study, we wish to stimulate future research on prior knowledge in knowledge acquisition processes. Therefore, we propose three suggestions. First, although there is a plethora of studies investigating knowledge as a dependent variable, there is much less research considering knowledge as an independent variable triggering non-cognitive learning processes. The field of motivational research in educational contexts is immense, but only a small proportion has considered assessing prior knowledge, motivation, and learning outcomes. For example, the construct academic self-concept stimulates lots of research, but only eight studies could be included in our meta-analysis, providing correlations between prior knowledge, self-concept, and learning outcomes. Despite sufficient studies assessing knowledge on two occasions and motivation, only a few studies explicitly addressed mediation processes. We consider it fruitful to address the mediation hypothesis concerning motivational constructs in further studies.

Second, as noted in the limitations, the results of the current study do not provide causal explanations. Although longitudinal study designs provide valuable insights into the role of prior knowledge on knowledge acquisition processes, we encourage researchers to conduct experiments in which prior knowledge is manipulated to investigate the causal effects of prior knowledge on knowledge acquisition. These experiments require a pre-learning intervention given to the experimental group to form randomized high and low prior knowledge groups. Existing designs use naturally occurring differences in prior knowledge to form groups, which makes these studies quasi-experimental studies, often influenced by confounding variables.

As the third impulse for future research, we want to address additional mediating processes not covered in the present study. There may be other non-cognitive mediators explaining the relation between prior knowledge and learning outcomes. For example, learners with high prior knowledge and sophisticated epistemic beliefs provided arguments of better quantity, quality, and diversity on socioscientific issues (Baytelman et al., 2020). Furthermore, we did not analyze models with parallel mediators. It seems reasonable to assume that the motivational constructs we considered are also intercorrelated. This raises the question of how much of the explained variance is unique to a specific type of motivation. Cheung (2021) discusses methods to conduct meta-analyses with multiple mediators simultaneously. Additionally, motivation may be part of a series of mediators. In models of self-regulated learning, motivation only provides the impetus for behavior that ultimately leads to storing information in long-term memory, such as learning strategies (Panadero, 2017). Therefore, other factors might mediate the relation between motivation and learning outcomes.

6.6 Conclusion

This study is the first meta-analysis to investigate explanations for interindividual stability in knowledge scores focusing on the role of motivation. We considered popular motivational constructs for our analyses: interest, self-concept, self-efficacy, intrinsic motivation, and extrinsic motivation. Prior knowledge assessed via objective knowledge tests was found to significantly predict each type of motivation at a later measurement point. We replicated the finding from Simonsmeier et al. (2021) regarding interindividual stability in knowledge scores and found that motivation partially accounts for it. However, the effects were highly heterogeneous, which hampers the prediction of future studies. Our moderator analyses revealed that the magnitude of effect sizes depended on methodological aspects and features of knowledge and motivation. The current meta-analysis proposed starting points for future research of great potential. The effects of prior knowledge on other non-cognitive variables are not yet studied carefully enough. Moreover, there may be more process variables explaining the relations between prior knowledge and learning outcomes as well as learning gains.

7 Study 3: Interventions to Foster Knowledge Transfer in School Students: A Meta-Analysis

7.1 Abstract

School students would ideally apply their acquired knowledge to new situations and problems. However, such knowledge transfer rarely occurs. Researchers have developed instructional interventions to promote transfer in students; but it is unknown to what extent and with which types of instructional methods transfer can be fostered best. This meta-analysis synthesizes studies evaluating how strongly instructional methods in transfer interventions foster knowledge transfer compared to control interventions. Herefore, we categorized instructional methods according to the learner's cognitive engagement and transfer processes. Additionally, we included information on the transfer distance based on the Taxonomy for Far Transfer (Barnett & Ceci, 2002). This meta-analysis used Hedges' g as effect size to evaluate the effect of transfer interventions on knowledge transfer in a random-effects meta-analysis with robust variance estimation. We predicted a medium total effect compared to control interventions as the main hypothesis. Aspects of the instructional methods and the transfer distance were included as moderators. Exploratory moderators were other characteristics of the intervention and the samples. The main hypothesis was confirmed: Transfer interventions were more successful in promoting knowledge transfer than control interventions across the included studies ($g^+ = 0.285$). All instructional methods were equally effective. However, transfer interventions were more effective for younger students and lower school levels. Also, the intervention setting in the field and successful implementation checks positively influenced the results. Besides several limitations and strengths, the results of this meta-analysis have important implications for educational research and practice.

7.2 Introduction

7.2.1 Knowledge Transfer

Knowledge transfer, or short *transfer*, is the learner's ability to apply learned content to similar or dissimilar tasks in new situations or contexts (Schwartz et al., 2008). It answers how learning in one task or situation impacts the responses in other tasks or situations (Adams, 1987; Anderson, 1983b; McGeoch, 1942; Perkins & Salomon, 1992). These responses include changes in performance because of prior performance in a different task (Gick & Holyoak, 1987). In other words, transfer contains the transmission of knowledge to new content areas as well as new physical, temporal, or functional contexts (Barnett & Ceci, 2002).

Thus, the more learners transfer their knowledge, the broader the range of new problems they can solve. Traditionally, this is a crucial goal at school. Besides the school context, transfer of trainings for adults at the workplace has been researched often (e.g., Blume et al., 2010; Cheng & Hampson, 2008; Ford & Weissbein, 1997). Studies in this research area answer questions on the causes of successful transfer in work contexts (Baldwin & Ford, 1988; Blume et al., 2010; Velada et al., 2007), factors that facilitate the transfer process (Grossman & Salas, 2011; Tannenbaum & Yukl, 1992) or who is most responsible for maximizing transfer in trainings (Liu et al., 2008; Subedi, 2004). Accordingly, fostering transfer has been one of the central goals of education in the last centuries (Butterfield & Nelson, 1989). In contrast to the research area of adult education, trainings to facilitate knowledge and their impact on transfer are less explicitly and broadly researched for school students.

Although knowledge transfer is a central goal of education, simultaneously, it has been one of the biggest challenges in the research of learning and teaching (Ormrod, 2012). Transfer occurs less than researchers expect, and learners hope (Haskell, 2001). Learners show little transfer because of inert knowledge, sole context-specific activation of knowledge, and the focus on the situations' surface characteristics (e.g., Chi & VanLehn, 2012). Teachers often must revise content because students do not recognize concepts and principles from their previously learned material (Kaminske et al., 2020).

Previous meta-analyses only partly assessed the field of knowledge transfer in terms of knowledge domains (e.g., Rittle-Johnson et al., 2017; Scherer et al., 2019) or instructional methods (e.g., Pai et al., 2015; Pan & Rickard, 2018). Rittle-Johnson and colleagues (2017) examined the influence of self-explanations on learning and transfer in mathematics. They formulated several guidelines for mathematics educators, including, for example, the design of explanations. As a second example, learning computer programming positively transferred to mathematical skills (Scherer et al., 2019). Both meta-analyses provide insights into the effectiveness of instructional methods in specific knowledge domains. On the other hand, the effectiveness of specific methods has also been meta-analytically aggregated. Examples are the influence of small-group learning (Pai et al., 2015) and test-enhanced learning (Pan & Rickard, 2018) on knowledge transfer. These studies provide first insights into the possible influences of specific instructional methods on knowledge transfer. However, these meta-analyses only comprised aspects of the broad domain of knowledge transfer.

7.2.1.1 Transfer Theories. There have been many theories on knowledge transfer in the last century. Reviews on these transfer theories usually start with the distinction between the theory of identical elements (Thorndike, 1924; Woodworth & Thorndike, 1901) and formal discipline (Judd, 1908; Marton, 2006; Royer, 1979). In particular, the two theories differ in the importance of similar features of the learning and transfer situation. Later theories concentrated more on cognitive features, for example, the similarity of productions (ACT* theory; Anderson, 1993), the importance of analogies (analogical transfer; Gentner & Colhoun, 2010), the differentiation between high and low road transfer (Salomon & Perkins, 1989), and perception-action processes (Day & Goldstone, 2012). On the other hand, theories that purported the importance of social and situational aspects are based on Lave (1988) and her critique of behavioral transfer theories. These sociocultural theories focused on affordances in situations

(Greeno et al., 1993), the framing of contexts (Engle et al., 2012), transfer as the preparation for future learning (Bransford & Schwartz, 1999), and the concentration on the learner's perspective (actor-oriented transfer; Lobato, 2003), among other aspects.

While researchers agree with the general existence of knowledge transfer, the theoretical and practical foundations differ considerably between theories. This makes it difficult to operationalize transfer and attribute success in empirical research to exact theoretical viewpoints. For this meta-analysis, we tried to systematize existing transfer theories based on two underlying cognitive processes in transfer, *Abstraction* and *Concretization*.

7.2.1.2 Abstraction and Concretization. Several theories on knowledge transfer focus on cognitive processes. They base on the assumption that the probability of retrieving relevant prior learning during search processes in memory determines successful knowledge transfer (Royer, 1979). Numerous cognitive mechanisms influence knowledge transfer in this search, for example, the encoding of representations of current situations or general prior learning (Butterfield & Nelson, 1989). Two other opposing cognitive processes are particularly relevant to examine transfer: the *Abstraction* and the *Concretization* of knowledge.

Abstraction describes the process of reducing details to subsume concrete examples under broad categories to generate less situation-specific knowledge (e.g., Goldstone & Sakamoto, 2003; Kaminski et al., 2008). Presenting content in an abstract and generic form facilitates knowledge transfer (Sloutsky et al., 2005) by not referencing concrete objects or contexts (Schalk et al., 2016). Central theories on transfer purport the importance of cognitive processes that can be associated with *Abstraction*: One of them is identifying identical elements in the source and the target domain so that knowledge acquired in one domain can be applied in a second domain (Thorndike, 1924). These identical elements were said to be stimulus-response pairs (Thorndike & Woodworth, 1901) or, later, shared productions (Singley & Anderson, 1989). Another relevant process connected to *Abstraction* is called analogical reasoning. Analogical reasoning aids learners in mapping systems of relations between knowledge elements (Mayer & Wittrock, 1996; Singley & Anderson, 1989; Vosniadou & Ortony, 1989). Through invoking analogies, learners can identify similar features and relationships between cases (Alfieri et al., 2013).

The *Abstraction* of learned contents from the original learning context is a crucial process in transferring knowledge. Hence, instructors should focus on abstracting contents, for example, by teaching knowledge of intermediate generality (Thomas et al., 1992). Also, reducing concrete topics improves a person's knowledge transfer to dissimilar situations (Day & Goldstone, 2012). Especially for young children, teaching more abstract tasks instead of concrete ones seems to be beneficial (DeLoache, 1991, 1995).

On the other side, the importance of *Concretization* is purported by sociocultural theories. These theories focus on the social and physical aspects of the learning situation (e.g., Suchman, 1987). According to this, learning is situated and cannot be separated from the social context (e.g., Greeno et al., 1993; Guberman & Greenfield, 1991; Roth & Jornet, 2013). Starting in the late 1980ies, Lave (1988) criticized that cognitive processes of transfer are described without their realistic context in previous transfer theories. Therefore, the theory of situated learning includes learning and cognition as situated processes based on activities in social contexts (Greeno et al., 1993). Moreover, learning is generally seen as a "process of enculturation or individual participation in socially organised practices" (Hennessy, 1993, p. 2). Transfer is realized through peripheral participation, the procedure of taking over behavioral and cognitive patterns (Brown et al., 1989; Lave & Wenger, 1991). Therefore, problem-solving and its transfer are bound to social practice (Crook, 1991) and the learning context. Other theories have been brought forth in this area: The perspectives include transfer as a preparation for future learning (e.g., Belenky & Nokes-Malach, 2012; Bransford & Schwartz, 1999; Schwartz & Martin, 2004; Stratton, 2020) and the concept of actor-oriented transfer (e.g., Lobato, 2003, 2006, 2012; Lobato & Siebert, 2002). Both criticize the traditional framing of transfer as the result of a transfer assessment determined by experts.

By focusing instructions on the *Concretization* of knowledge, transfer is enhanced. The use of concrete materials promotes the comprehensibility of tasks (Bruner, 1966). Learners concentrate on concrete examples rather than abstract verbal instructions (LeFevre & Dixon, 1986). Additionally, students who learned with concrete physical bills were less likely to make conceptual errors in mathematical operations (McNeil et al., 2009).

Instructional methods in interventions differ in their promotion of either the process of *Abstraction* or *Concretization* of knowledge while learning. Therefore, most instructional methods in interventions can be assigned to either of these two categories. This makes it possible to investigate the relative importance of each of the proposed processes for knowledge transfer based on the methods in an intervention.

7.2.2 Instructional Methods

Research on the effectiveness of instructional methods has its roots in the domain of instructional-design (e.g., An, 2021; Gagné, 1965; Glaser, 1962; Reiser, 2018). Theories on instructional-design generally contain guiding principles on how to help learners learn (Reigeluth, 1999; Reigeluth & An, 2020). More specifically, these theories focus on the design and methods of instructional interventions in varying detail (e.g., Cronbach & Suppes, 1969; Perkins, 1992; Reigeluth, 1983; Reigeluth & Merrill, 1979).

Instructional methods are “anything that is done purposely to facilitate learning or human development” (Reigeluth & Carr-Chellman, 2009, p. 21). They are componential and probabilistic (Reigeluth, 1999). Componential means that instructional methods can be defined on several levels. Each level contains components that may differ depending on the application criteria in interventions. Instructional methods are also probabilistic because applying methods does not guarantee the intended outcome; they only increase the probability of it (Reigeluth, 1999). This probability of success depends on the used method components (Reigeluth, 1999;

Reigeluth & An, 2020) as well as on the situation, which comprises aspects of the learner, the content, and the context (Honebein, 2022; Honebein & Reigeluth, 2020).

There are many instructional methods (e.g., Honebein & Reigeluth, 2020; Reigeluth, 1999; Treaugust, 2014) that differ in several aspects, for example, the scope or the generality (Reigeluth & An, 2020; Reigeluth & Carr-Chellman, 2009). Also, descriptions of instructional methods vary in their broadness. Reigeluth and Merrill (1979) proposed three classes of instructional variables. According to this, in this meta-analysis, we concentrate on methods that were classified as delivery strategy variables: These are methods for conveying the instruction to the learners (Reigeluth, 1983) and that are used to carry out the instructional process (Reigeluth & Merrill, 1979). Such instructional methods can further be distinguished in their fidelity of representation, interactiveness, special capabilities, motivational effects, and cost (Reigeluth & Merrill, 1979). Other differentiations include the amount of control instructors have (e.g., Treaugust, 2014) or the student's cognitive processing (e.g., Sternberg, 1981; Tobias, 1982). From the perspective of the instructional-design theory, it is plausible that instructional methods differ in their promotion of knowledge transfer (e.g., Clark & Voogel, 1985; Fries et al., 2021; Van Merriënboer & Kirschner, 2017).

7.2.2.1 Categorization of Instructional Methods. There are several frameworks with categorizations for instructional methods that can be used in the context of learning. However, most of these frameworks do not explicitly focus on the impact of instructional methods on knowledge transfer. The ICAP framework (Interactive-Constructive-Active-Passive framework; Chi, 2009; Chi & Wylie, 2014) incorporates students' overt learning behaviors as the basis of their cognitive engagement as reactions to instructional methods and their relations to knowledge transfer as a cognitive outcome. It incorporates a taxonomy with four hierarchical modes of students' cognitive engagement. The modes of passive, active, constructive, and interactive cognitive engagement have different predicted cognitive outcomes. Whereas passive instructional methods only foster the recall of learned material, the other three modes promote the application as well as the transfer of knowledge to new contexts.

The four modes are independent of the instruction itself. This means that various modes of cognitive engagement can occur while learning, regardless of the methods used (Chi & Wylie, 2014). Nevertheless, instructional methods can be assigned to the four modes based on the expected typical learner's cognitive engagement. The ICAP framework, initially proposed as a theory for interpreting learning research findings (Chi, 2009), has been applied to instructional designs in several studies (e.g., Roscoe et al., 2014). Therefore, it is well suited to categorize the instructional methods of interventions to foster knowledge transfer in school students (Wiggins et al., 2017). Instructional methods that elicit active, constructive, or interactive cognitive engagement in students should enable them to transfer learned content.

7.2.2.2 Combining Cognitive Engagement with Abstraction and Concretization. To structure instructional methods, we combined methods according to the learner's cognitive engagement from the ICAP framework (Chi & Wylie, 2014) and the cognitive processes of *Abstraction* and *Concretization* to form a theoretical matrix (see Table 10). Some moderating influences can be expected: We anticipated interactive instructional methods to be the most effective, followed by constructive, active, and passive methods, respectively. Additionally, we assumed that instructional methods fostering the *Abstraction* of knowledge promote transfer better than methods fostering *Concretization*. To assess interactions between the two theoretical models, the combination of concrete passive and active methods as well as the combination of abstract constructive and interactive methods, were expected to be more effective in promoting transfer than their counterparts, concrete constructive and interactive methods, as well as abstract passive and active methods.

Table 10

Schematic Representation of the Theoretical Matrix for Instructional Methods

	Abstraction	Concretization
Passive		
Active		
Constructive		
Interactive		

Note. Matrix of ICAP framework's modes of cognitive engagement (Passive, Active, Constructive, and Interactive) and the cognitive processes Abstraction and Concretization.

7.2.3 Moderators of the Relation Between Interventions and Knowledge Transfer

Moderators might influence the relation between interventions and transfer. Previous findings purported the importance of the learner's cognitive engagement (e.g., Chi & Wylie, 2014) and cognitive processes (e.g., Day & Goldstone, 2012). Furthermore, transfer assessments depended on the transfer distance in prior research (e.g., Perkins & Salomon, 1992). Additionally, we included several exploratory moderators in our analyses that have not yet been inspected closely in this context.

First, we expect that the promotion of interventions on transfer differs according to the instructional methods in the beforementioned theoretical matrix with modes of cognitive engagement and transfer processes. Interactive methods should foster transfer best, followed by constructive, active, and passive methods. Adding to this, we anticipate methods that include the process of *Abstraction* to promote transfer better than methods of *Concretization*.

Knowledge transfer also differs based on the assessed type of transfer. Researchers distinguish several types of transfer (e.g., Royer, 1979), for example, near and far transfer (e.g., Ormrod, 2012). Transfer is called near when knowledge is transferred to a primarily similar context. This type of transfer is found relatively often in empirical studies (Perkins & Salomon, 1992). Far transfer, in which knowledge is transferred to a dissimilar context, is rarer (Haskell, 2001). To further examine the transfer distance measured in the primary studies, we used the

Taxonomy for Far Transfer (Barnett & Ceci, 2002): The taxonomy allows categorizing the transfer distance depending on the similarity of the learning and application contexts. Context similarity is divided into six context factors: *Knowledge Domain, Physical Context*, *Temporal Context*, *Functional Context*, *Social Context*, and *Modality* (Barnett & Ceci, 2002). We expect that interventions better promote near transfer than far transfer and that the effectiveness of instructional methods differs along the six context dimensions.

Other factors can be expected to influence the effectiveness of transfer interventions. Examples are different subpopulations, for example, the age and school level, and characteristics of the interventions. Additionally, differences between knowledge types and methodological characteristics could be substantial. These factors were inspected exploratively.

7.2.4 The Current Study

In this meta-analysis, we investigated the average effect of interventions (*transfer interventions*) on knowledge transfer in school students compared to control groups and comparison interventions (*control interventions*). Knowledge transfer is a broad field of research. Other meta-analyses only partly assessed this field in terms of restrictions on knowledge domains (e.g., Rittle-Johnson et al., 2017; Scherer et al., 2019), or on included instructional methods (e.g., Pai et al., 2015; Pan & Rickard, 2018). These narrower meta-analyses leave a research gap open, which currently hinders further insights into this research domain. The current meta-analysis closes this gap by including primary studies from various knowledge domains and instructional methods on different transfer distances. This leads to the question: How effective are transfer interventions generally in promoting knowledge transfer compared to control interventions? To answer this question, we formulated the main hypothesis (H1): Instructional transfer interventions have a medium effect on knowledge transfer averaged over studies and instructional methods compared to control interventions.

Primary studies seldom used operationalizations provided by transfer theories. Therefore, we used empirical evidence from classical transfer studies and intervention studies focusing on knowledge acquisition in educational contexts that assessed knowledge transfer as a side result.

Other variables might influence the effectiveness of interventions on knowledge transfer. To examine this heterogeneity, several moderators are considered: First, a theoretical matrix comprising the modes of cognitive engagement from the ICAP framework (Chi & Wylie, 2014) as well as the cognitive processes *Abstraction* and *Concretization* were used to categorize instructional methods. Adding to this, we incorporated the differentiation in near and far transfer on the six context factors of the Taxonomy for Far Transfer (Barnett & Ceci, 2002) as moderators. Several additional exploratory moderators were also included in this meta-analysis.

The meta-analysis investigates how instructional methods influence the effectiveness of interventions on knowledge transfer. We formulated a theoretical two-dimensional matrix to sort the instructional methods: It combined the cognitive processes, *Abstraction* and *Concretization* of knowledge, and the four modes of cognitive engagement. The allocation to this categorization was based on the typical application of the instructional methods in primary studies. We formulated three moderator hypotheses based on this model: (H2) Interactive instructional methods are most effective, followed by constructive, active, and passive methods. (H3) Instructional methods that foster the *Abstraction* of knowledge promote transfer better than methods that foster *Concretization*. (H4) The combination of concrete passive and active methods as well as the combination of abstract constructive and interactive methods are more effective in fostering transfer than their counterparts in the theoretical matrix.

In which way does the transfer distance impact the effectiveness of interventions on knowledge transfer? To answer this question, we examined the transfer distance measured in the primary studies with the Taxonomy for Far Transfer (Barnett & Ceci, 2002). So, we could inspect which intervention fostered near or far transfer on which dimension of the six context factors. To assess this moderator, we formulated two hypotheses: (H5a) Interventions better promote near transfer than far transfer, and (H5b) the effectiveness of instructional methods differs along the six transfer dimensions. To test hypothesis H5a, we compared interventions based on the transfer distance of each of the six context factors. H5b is tested by inspecting the interaction between instructional methods and the six context factors.

We further examined several exploratory moderators to determine the generalizability of the advantage of transfer interventions compared to control interventions. We also analyzed sample characteristics as well as characteristics of the intervention and the methodology.

7.3 Methods

7.3.1 Preregistration

This meta-analysis was preregistered (DOI: http://dx.doi.org/10.23668/psycharchives.10015). Since then, the following changes occurred: We concretized the main hypothesis in its wording regarding the comparison of experimental and control groups. Additionally, we conducted a latent class analysis to reduce the number of instructional methods that would have been inspected as individual moderators. Therefore, we also included latent classes in the moderator analysis besides analyzing particular instructional methods.

7.3.2 Literature Search

Figure 28 summarizes the search and inclusion process for this meta-analysis. Data collection included a systematic search conducted in April 2022 in two databases (PsycINFO and ERIC) with the following search terms: (knowledge* or concept* or principle* or procedure* or skill*) and transfer and (train* or instruct* or facilitat* or interven* or lesson* or learning environment*). This search string covers terms for knowledge, transfer, and synonyms for interventions. Additionally, the search limits were set to non-disordered population and childhood as well as adolescence. This literature search yielded 1663 results in PsycINFO. Due to search restrictions in ERIC, other operationalizations of the limitations were used: The limit was set to educational settings from early childhood education to secondary education (early childhood education or elementary education or secondary education or grade

1 or grade 2 or grade 3 or grade 4 or grade 5 or grade 6 or grade 7 or grade 8 or grade 9 or grade 10 or grade 11 or grade 12 or high school equivalency programs or high schools or intermediate grades or junior high schools or kindergarten or middle schools or preschool education or primary education or secondary education). This search yielded 1560 results. After the deletion of duplicates, 2984 search results remained. An additional explorative search based on cited literature in primary studies, reviews, and Google Scholar produced 122 results.

7.3.3 Inclusion of Studies

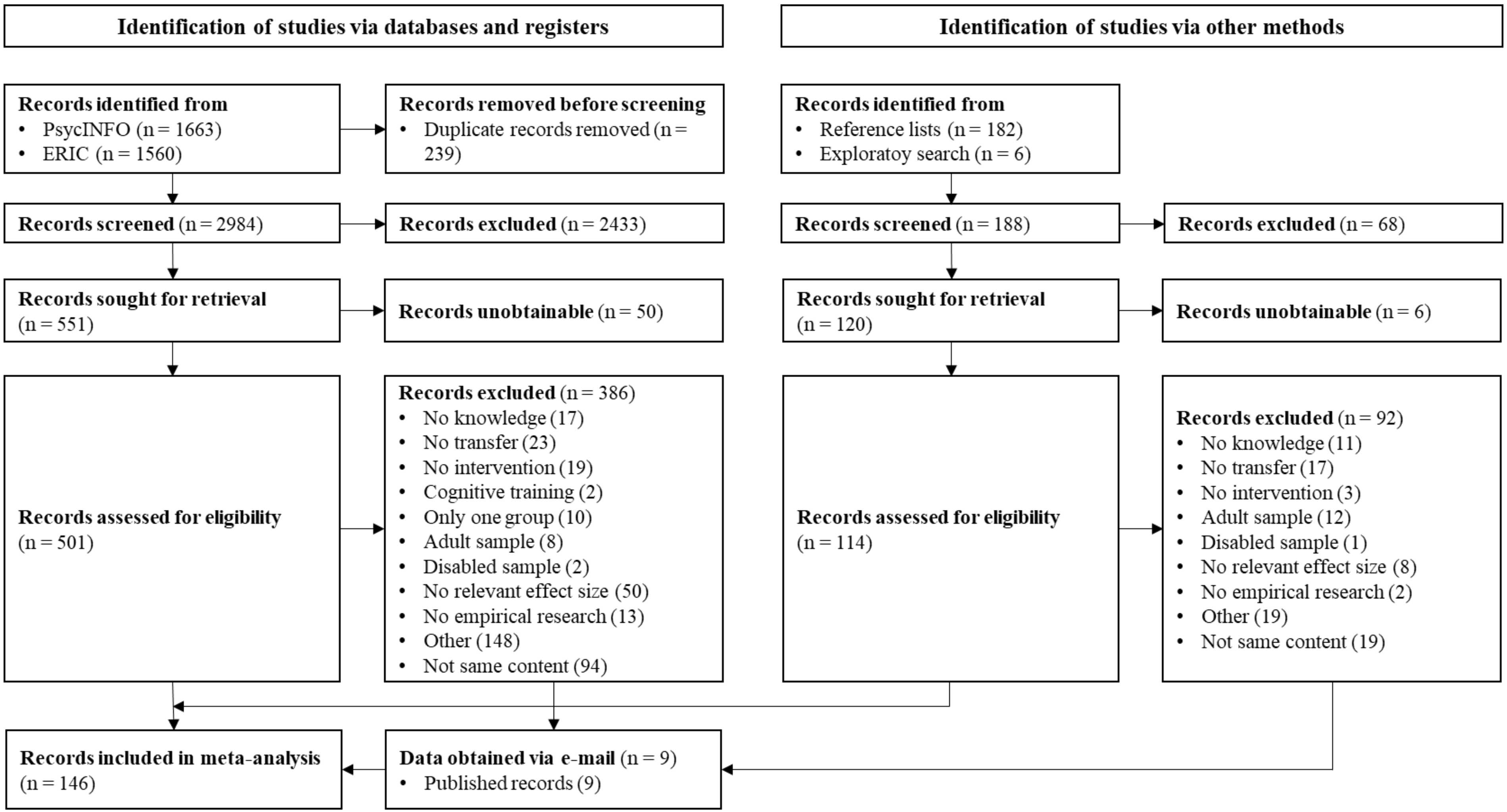
We used the following inclusion criteria to screen titles, abstracts, and full texts: (a) The study included at least one measurement point after an instructional intervention that provided learning of domain-specific knowledge. (b) The knowledge measured after the instructional intervention must be identifiable as transfer of the initially learned contents. It comprised only objective quantitative measures of domain-specific knowledge. Measures of meta-cognitive knowledge, executive functions, intelligence, basic skills, and self-assessments were excluded. (c) The study included an instructional intervention with an experimental or a quasi-experimental design and at least one group that could be identified as a control or comparison group. Both groups learned the same content in the intervention. (d) The study included participants whose mean age was 18 or younger. (e) The study provided the information necessary to compute standardized effect sizes. Studies that did not fit these inclusion criteria were excluded.

Further exclusion of studies was conducted for the following reasons: We excluded studies in which instructional methods did not differ between groups (e.g., Schneider et al., 2015) and studies that did not provide sufficient descriptions of the interventions (e.g., Toll & Van Luit, 2014) or transfer measurements (e.g., Peled & Segalis, 2005).

We determined study eligibility based on the inclusion criteria in two steps: First, the first author and four research assistants screened titles and abstracts of the search hits. The

Figure 28

Flow Chart for the Literature Search and Inclusion Process (Study 3)



Note. Flow diagram of the study search and inclusion process (adapted from Page, McKenzie, et al., 2021; Page, Moher, et al., 2021). The process on the left is based on records identified by the systematic search, and the process on the right is based on the exploratory search.

studies were either excluded with reasons or retained to examine the full texts. The research assistants double-coded 100 abstracts independently after receiving coder training with the first author. The individual agreements were between 80% and 84%. Disagreements were resolved by discussion. Two research assistants independently screened each remaining abstract, and the first author solved disagreements. In the second step, the full texts were screened and excluded if they failed to meet the inclusion criteria. Based on coder training, the first author and a research assistant independently coded 50 studies for eligibility. The mean interrater agreement for this step was 86%, and disagreements were again resolved by discussion. The first author then screened the remaining full texts.

After contacting authors of articles that did not provide the relevant information, data from nine studies was provided via mail. Forty-nine studies could not be included because the relevant information was not provided in the study and by the authors. At this point, 811 effect sizes from 146 articles were included in this meta-analysis.

7.3.4 Data Coding

Data coding was based on a coding manual that entailed predefined coding rules. This manual was used for coder training and coding of the articles. The following information was gathered for all included studies (see Table 11): We coded several study characteristics, such as the authors, the year of publication, and whether the study has been published in a peer-reviewed journal or as grey literature. Also, we coded whether the instructional intervention was explicitly based on a theory of knowledge transfer. Additionally, the total *N* of both groups in the study, the study design and assignment of the participants to groups, the setting, as well as the country and the continent were coded. We included the number and duration of sessions and the days between the last session and the transfer assessment for each group.

Characteristics of the sample were coded for the experimental and the comparison group. These are the *N* of the group, the school level, the class, the gender of participants, the age group, and the mean age. For cases when the intervention involved more than one class level, we used the mean of the class levels if these were in a maximum range of three classes. Otherwise, the class was coded as missing information. The intervention characteristics included the group size of the intervention, information about the trainer person and sessions (e.g., use of technology), as well as information on implementation checks.

Table 11

Description of the Included Moderators (Study 3)

Moderator	**Description**
Intervention characteristics	
Setting	To test whether the influence of interventions on knowledge transfer differs between studies using different settings, we coded two settings: *field* and *lab*.
Mean number of sessions	We coded the mean number of sessions for the experimental and comparison groups.
Mean duration of sessions	We coded the mean duration in minutes of sessions for the experimental and comparison groups.
Mean days between intervention and transfer test	We coded the mean number of days between the intervention and the transfer test for the experimental and comparison groups.
Broad content domain	The effectiveness of interventions on knowledge transfer might differ based on the broad content domain of the training. We coded seven broad content domains: *STEM, language, social sciences, health sciences, humanities, movement,* and *other*.
Content domain	To test whether the influence of interventions on knowledge transfer differs between studies analyzing interventions in different content domains, we coded 13 content domains: *Mathematics, physics, chemistry, biology, geosciences, engineering, first language, foreign language, history, sports, motion, other,* and *several*.
Group size of intervention	We coded the group size of the intervention as four levels: *single, small group, whole class,* and *several*.
Trainer person	We distinguished between interventions with different trainer persons: *researcher, teacher, other,* and *mixed*.
Technology-enhanced learning	To test whether the influence of interventions on knowledge transfer differs between studies with and without technology-enhanced learning, we coded the levels *yes* and *no*.
Intelligent tutor	To test whether the influence of interventions on knowledge transfer differs between studies with and without an intelligent tutor, we coded the levels as *yes* and *no*.
Intention to transfer	To test whether the influence of interventions on knowledge transfer differs between studies with and without the intention to transfer, we coded the levels *yes* and *no*.

Table 11 (continued)

Moderator	Description
Instructional methods	
Instructional methods	We coded 27 instructional methods with *yes* or *no*: *advance organizer, annotated examples, collaborative learning, comparison of annotated examples, comparison of concepts, concrete models, concreteness fading, demonstration, drill, examples of concepts, explanation, feedback, generate drawings, imitation, interleaved practice, mnemonic systems, movement, multiple representations, paraphrasing, problem posing, realistic problems, rehearsal, scaffolding, self-explanation, variability of practice and tasks, verbal or written instruction,* and *visualizations.*
Sample characteristics	
Age	We coded the sample mean age of the learners (in years) for the experimental and comparison group.
Age group	We distinguished between different age groups for the experimental and comparison group. We coded the following levels: *under 6 years, 6 to 11 years, 12 to 18 years,* and *several.*
School level	To test whether the influence of interventions on knowledge transfer differs between studies including different school levels for the experimental and comparison group, we coded the levels *kindergarten and preschool, primary school, secondary school,* and *several.*
Class	We coded the sample mean class of the learners for the experimental and comparison group.
Gender	We distinguished three levels for the experimental and comparison group: *mixed, female,* and *male.*
Transfer characteristics	
Knowledge domain	We coded the context factor knowledge domain as *near* or *far.*
Physical context	We coded the context factor physical context as *near* or *far.*
Temporal context	We coded the context factor temporal context as *near* or *far.*
Functional context	We coded the context factor functional context as *near* or *far.*
Social context	We coded the context factor social context as *near* or *far.*
Modality	We coded the context factor modality as *near* or *far.*
Methodological study characteristics	
Grey literature	We coded whether the primary study is grey literature with *yes* or *no*.
Study design	We coded whether the primary study was a *randomized controlled trial* or *quasi-experimental*.
Setting pre-test	We distinguished between the different settings of the pre-test with two levels: *lab* and *field*.
Setting post-test	We coded the setting of the post-test with two levels: *lab* and *field*.
Number of items	We coded the mean number of items in the transfer measurement.
Knowledge type	To test whether the transferred knowledge type influences the relation between interventions and knowledge transfer, we coded three levels: *declarative, procedural,* and *mixed*.

Table 11 (continued)

Moderator	Description
Specific knowledge type	To test whether the specific transferred knowledge type influences the relation between intervention and knowledge transfer, we coded five levels: *facts, concepts, cognitive skill, motor skill,* and *mixed*.
Response format	To test whether the response format of the transfer assessment influences the relation between interventions and knowledge transfer, we coded three levels: *open, multiple-choice,* and *mixed*.
Implementation check	We coded three levels for the implementation check: *yes & successful, yes & not successful,* and *no*.
Continent	We coded the continent on which the primary study took place.

Note. Overview of included moderators and the coded moderator levels. STEM – Science, technology, engineering, mathematics.

Additionally, we coded instructional methods that were present in the primary studies. We included 25 pre-defined instructional methods based on previous research on knowledge transfer (e.g., Jelsma et al., 1990). These were *Advance Organizer*, *Annotated Examples*, *Collaborative Learning*, *Comparison of Annotated Examples*, *Comparison of Concepts*, *Concrete Models*, *Demonstration*, *Drill*, *Explanations*, *Feedback*, *Generate Drawings*, *Imitation*, *Interleaved Practice*, *Mnemonic Systems*, *Multiple Representations*, *Paraphrasing*, *Problem Posing*, *Realistic Problems*, *Rehearsal*, *Scaffolding*, *Self-Explanations*, *Variability of Practice*, *Verbal or Written Instructions* and *Visualizations*. In the coding process, we added the instructional methods *Concreteness Fading* and *Movement* based on primary studies. Other methods, coded in the category *Other*, were left out for further analyses because of the scarcity of those methods and the variability of this category. Each instructional method was coded based on definitions provided by earlier research (see Table 12). We only coded instructional methods that the experimental group received but not the comparison group. Instructional methods used in both the treatment and the comparison group were omitted.

Table 12

Overview of Instructional Methods

Instructional Method	Description
Advance Organizer	"In all cases I define advance organizers as introductory material at a higher level of abstraction, generality, and inclusiveness than the learning passage itself, and an overview as a summary presentation of the principal ideas in a passage that is *not necessarily* written at a higher level of abstraction, generality, and inclusiveness, but achieves its effect largely by the simple omission of specific detail (Ausubel, 1963, 1968)" (Ausubel, 1978, p. 252).
Annotated Examples	"As instructional devices, they [worked examples] typically include a problem statement and a procedure for solving the problem; together, these are meant to show how other similar problems might be solved. In a sense, they provide an expert's problem-solving model for the learner to study and emulate" (Atkinson et al., 2000, pp. 181-182).
Collaborative Learning	"CL [Collaborative learning] is an educational approach to teaching and learning that involves groups of learners working together to solve a problem, complete a task, or create a product" (Laal & Laal, 2012, p. 491).
Comparison of Annotated Examples	"Comparison is a general learning process that can promote deep relational learning and the development of theory-level explanations (Forbus, 2001; Gentner, 1983, 2003). According to structure-mapping theory, comparison acts to highlight commonalities, particularly relational commonalities that may not have been noticed prior to comparison (Gentner, 1983, 2003)" (Gentner, 2005, p. 251).
Comparison of Concepts	"Comparison is a general learning process that can promote deep relational learning and the development of theory-level explanations (Forbus, 2001; Gentner, 1983, 2003). According to structure-mapping theory, comparison acts to highlight commonalities, particularly relational commonalities that may not have been noticed prior to comparison (Gentner, 1983, 2003)" (Gentner, 2005, p. 251).
Concrete Models	"A *concrete model* represents a mathematical idea by means of three-dimensional objects" (Fennema, 1972, p. 635). "Concrete materials, which include physical, virtual, and pictorial objects, are widely used in Western classrooms (Bryan et al. 2007), and this practice has support in both psychology and education (e.g., Bruner 1966; Piaget 1970)" (Fyfe et al., 2014, p. 10).
Concreteness Fading	"Specifically, we argue for an approach that begins with concrete materials and gradually and explicitly fades toward more abstract ones. This concreteness fading technique exploits the continuum from concreteness to abstractness and allows learners to initially benefit from the grounded, concrete context while still encouraging them to generalize beyond it" (Fyfe et al., 2014, p. 10).
Demonstration	"This [demonstration] involves intentionally showing somebody else how to perform a task or how to solve a problem" (Gärdenfors, 2017, p. 2).
Drill	"Drill, then, means the 'formation of habits through regular practice of stereotyped exercises'" (Schofield, 2013, p. 45).
Examples of Concepts	"Concerning instruction of declarative concepts, a common feature of textbooks and of lectures is to introduce a declarative concept by presenting the definition and then to elaborate further by describing concrete examples of how that concept can be applied in one or more real-world situations (hereafter referred to as illustrative examples)" (Rawson et al., 2015, p. 484).

Table 12 (continued)

Instructional Method	Description
Explanations	"Explanations refer to how or why a phenomenon occurs (Chin & Brown, 2000)" (McNeill & Krajcik, 2008, p. 54). "An "explanation" in science should make sense of a phenomenon, based on other scientific facts, and establish the relationships, based on evidence and logical reasoning (Berland and Reiser 2008; Osborne and Patterson 2011)" (Yang & Wang, 2014, p. 533).
Feedback	"[…] feedback is conceptualized as information provided by an agent (e.g., teacher, peer, book, parent, self, experience) regarding aspects of one's performance or understanding" (Hattie & Timperley, 2007, p. 81). "*Feedback* is a key element in formative assessment, and is usually defined in terms of information about how successfully something has been or is being done" (Sadler, 1989, p. 120).
Generate Drawings	"In sum, the learner-generated drawings […] are defined as pictorial representations (a) that are intentionally constructed to meet a learning goal, (b) that are meant to depict represented objects accurately and, (c) for which the learner is primarily responsible for construction and/or final appearance" (Van Meter & Garner, 2005, p. 290).
Imitation	"Tomasello (1999) distinguishes between learning by *emulation* […] and learning by *imitation*, where the learner observes the sequence of the model's actions and tries to perform the same actions (process-oriented learning) (see also Zentall, 2001; Tehrani and Riede, 2008)" (Gärdenfors, 2017, p. 2).
Interleaved Practice	"If multiple kinds of skills must be learned, the opportunities to practice each skill may be ordered in two very different ways: blocked by type (e.g., aaabbbccc) or interleaved (e.g., abcbcacab)" (Rohrer & Pashler, 2010, p. 409).
Mnemonic Systems	"In this discussion a mnemonic device is considered to be a strategy for organizing and/or encoding information with the sole purpose of making it more memorable" (Bellezza, 1981, p.252).
Movement	"[…] non-verbal behavior - that is, any movement or position of the face and/or the body […]" (Ekman & Friesen, 1969, p. 49). → Non-verbal behavior by the participants that is prompted by instructors (e.g., tracing, gesturing, movements in sports)
Multiple Representations	"Early research on learning with MERs [multiple external representations] concentrated on the ways that presenting pictures alongside text could improve readers' memory for text comprehension (e.g., Levin, Anglin, & Carney, 1987). In the last two decades, the explosive increase in multi-media learning environments have widened the debate to include combinations of representations such as diagrams, equations, tables, text, graphs, animations, sound, video, and dynamic simulations" (Ainsworth, 2006, p. 187).
Paraphrasing	"[…] reiterating the information presented in the text in their own words such that no new information was added to the material in the form of an explanation" (Ainsworth & Burcham, 2007, p. 293).
Problem Posing	"Problem posing refers to both the generation of new problems and the re-formulation, of given problems. Thus, posing can occur before, during, or after the solution of a problem" (Silver, 1994, p. 19).
Realistic Problems	"[…] most real-world problems are ill defined. […] Ill-defined problems often require an individual to search outside sources to find relevant and potentially helpful information (Simon, 1978)" (Ormrod, 2012, pp. 423 - 424).

Table 12 (continued)

Instructional Method	Description
Rehearsal	"Rehearsal strategies emphasize repetition in various forms ranging from simply repeating the names of colors in the spectrum to underlining notes. In general, rehearsal strategies are designed to facilitate verbatim recall" (Weinstein, 1987, p. 592).
Scaffolding	"The first key feature that distinguishes scaffolding from other forms of instructional support is that it is temporary support that is provided as students are engaging with problems (Belland, 2014; Collins et al., 1989; Wood et al., 1976)" (Belland, 2017, p. 18). "Next, scaffolding needs to lead to skill gain such that students can function independently in the future (Belland, 2014; Pea, 2004; Wood et al., 1976)" (Belland, 2017, p. 18).
Self-Explanations	"Self-explanation is an instructional method in which learners are prompted to explain to themselves (either orally or in writing) as they study a lesson (such as a textbook chapter or a multimedia lesson)" (Johnson & Mayer, 2010, p. 1246).
Variability of Practice	"[…] the confrontation with a highly varied sequence of problems and/or solutions to those problems […]" (van Merriënboer, 2012, p.3390). "[…] Battig has expanded his earlier interpretation of intratask interference (Battig, 1972) to represent more general "contextual interference" including factors extraneous, as well as intrinsic, to the task being learned" (Shea & Morgan, 1979, p. 179).
Verbal or Written Instructions	Verbal instructions are "[…] explicit instructions [that] provide precise technical step-by-step instructions […]" (Meier et al., 2020, p. 2). A written instruction is "[…] a written description of the task […]" (LeFevre & Dixon, 1986, p. 1).
Visualizations	"Visualization is the ability, the process and the product of creation, interpretation, use of and reflection upon pictures, images, diagrams, in our minds, on paper or with technological tools, with the purpose of depicting and communicating information, thinking about and developing previously unknown ideas and advancing understandings" (Hershkowitz et al., 1989, p. 75).

Note. Overview of instructional methods included in the meta-analysis with the respective theoretical definition for coding the information.

We categorized the 27 instructional methods in the two-dimensional theoretical matrix (see Table 13). Eleven methods were identified as passive methods, whereas ten and five were classified as active and constructive, respectively. Only one method, *Collaborative Learning*, was inserted as an interactive method. Overall, we classified 11 methods as belonging to *Abstraction* and 16 methods as *Concretization*.

Table 13

Assignment of Included Instructional Methods to Theoretical Matrix

	Abstraction	Concretization
Passive	Advance Organizer Explanations Multiple Representations Scaffolding	Annotated Examples Concrete Models Demonstration Examples of Concepts Feedback Verbal or Written Instructions Visualizations
Active	Concreteness Fading Interleaved Practice Variability of Practice	Drill Imitation Mnemonic Systems Movement Paraphrasing Realistic Problems Rehearsal
Constructive	Comparison of Annotated Examples Comparison of Examples Self-Explanations	Generate Drawings Problem Posing
Interactive	Collaborative Learning	

Note. Matrix of ICAP framework's modes of cognitive engagement (Passive, Active, Constructive, and Interactive) and the cognitive processes Abstraction as well as Concretization with the assignment of instructional methods to the categories based on their usual use in interventions.

The coded characteristics of knowledge transfer comprised the differentiation in near and far transfer on the context factors *Knowledge Domain*, *Physical Context*, *Temporal Context*, *Functional Context*, *Social Context*, and *Modality* (Taxonomy for Far Transfer, Barnett & Ceci, 2002). The measurement characteristics contained the measurement of knowledge transfer at pre-test and post-test, each with the information about the broad content area, the more specific content area, the specific content, the knowledge type, the specific knowledge type, the number of items, the response format, the reliability of the measurement, and the group size of the assessment. The content of the instructional intervention was coded as the broad content area, the specific content area, and the specific content. Last, the means and standard deviations of both groups at pre-test and post-test, other effect size information given (e.g., Cohen's *d*) as well as the correlation between pre-test and post-test were coded as statistical information.

For most studies, the study design differentiated experimental and comparison groups. However, some studies included groups without such a differentiation: For those studies that did not provide a classified control or comparison group (e.g., Day & Cordón, 1993), we defined the group that received fewer instructional methods as the comparison group. The same approach was used for studies that involved several groups without a defined control group (e.g., Johnson et al., 2014).

In the third step, the first author and a research assistant double-coded the information of 50 effect sizes. We determined the agreement of categorical variables by percentage; the coding of continuous variables with interval scaling was correlated. The variables general and specific knowledge types were removed from the analyses because of poor agreement (60% for general knowledge type, 66% for specific knowledge type). This result can be attributed to the weak distinction between declarative and procedural knowledge in assessments outside the mathematics domain. The resulting mean interrater agreement for categorical variables was 94%, ranging from 70% to 100%. The mean correlation for variables with interval scaling was $r = .985$, ranging from $r = .856$ to $r = 1$. Again, disagreements were solved through discussion.

We contacted corresponding authors if the information required for the meta-analysis was not reported in a publication. Thirty-three authors were contacted to request additional information. Sixteen authors replied, of which eight supplied missing information (response rate: 68%). Moreover, a risk of bias assessment was conducted for all included primary studies. The questions for and interpretation of the Risk of Bias assessment were based on RoB2 (Sterne et al., 2019) and modified for our purposes (see Tables 22 and 23 in Appendix K).

7.3.5 Preparation of Effect Sizes

We used Hedges' *g* (g^+) as effect size and converted the given statistical information into it. These were mainly means and standard deviations and, in some cases, *F*-values and Cohen's *d*. For all studies that provided two measure points, we computed Hedges' *g* as the mean difference between post-test and pre-test scores, divided by the pooled sample standard deviations of the treatment and the comparison group (Morris, 2008). We computed Hedges' *g* based on the post-test score and the pooled sample standard deviations of the treatment and the comparison group for all studies with a post-test but no pre-test. We inspected the distribution of the effect sizes (e.g., skewness and kurtosis) before averaging them in the meta-analysis.

To correct measurement errors in the dependent variables, we used an attenuation correction based on the reliability provided in the studies. If no reliability was reported, we did not amend the effect size. The correction was done by dividing Cohen's *d* by the root value of the reported measurement's reliability (Schmidt & Hunter, 2015, p. 304; Wiernik et al., 2020).

We tested for outliers by inspecting Cook's values (Cook & Weisberg, 1982) using the *metafor* package (Viechtbauer, 2010). Cook's values, also called Cook's distance, measure the influence of effect sizes on the main result. This is done by removing each effect size from the model and summarizing the change this makes in the resulting model (Viechtbauer & Cheung, 2010). Through this process, outliers that have an excessive influence on the model and need further inspection can be identified. Rules of thumb promote several possible cut-off points. We removed effect sizes with a Cook's distance larger than four divided by the total number of effect sizes. This applied to 14 effect sizes from 13 studies (Britt & Aglinskas, 2002; Clarke et al., 1969; Durnin et al., 1997; Farkas, 2003; Kapur, 2010, 2012; Lamnina & Chase, 2021; McNeil et al., 2019; Ngu et al., 2009; Obiagu et al., 2020; Rau et al., 2015; Wagensveld et al., 2015; Williams & Carnine, 1981). These effect sizes were removed before the following analyses.

7.3.6 Statistical Analysis

7.3.6.1 Publication Bias. We tested for publication bias visually as well as statistically. For this, we used funnel plots and Egger regressions (Egger et al., 1997). These tests were performed using the *metafor* package (Viechtbauer, 2010) in R (R Core Team, 2022).

7.3.6.2 Meta-Analytic Integration. We expected heterogeneity in the included primary studies. Therefore, we conducted a random-effects meta-analysis (Cooper et al., 2018). To inspect differences between groups, we used Hedges' *g* as the effect size to evaluate the effect of transfer interventions on knowledge transfer compared to control interventions. We carried out the analysis with the *metafor* package (Viechtbauer, 2010) in R (R Core Team, 2022).

Most of the primary studies provided several effect sizes. These effect sizes were statistically dependent. Therefore, they violated a central assumption of meta-analytical models (e.g., Cheung, 2019). To handle these dependent effect sizes, we used robust variance estimation (Hedges et al., 2010) provided by the *clubSandwich* package (Pustejovsky & Pustejovsky, 2020) in R (R Core Team, 2022).

7.3.6.3 Moderator Analyses. We used meta-regressions with robust variance estimation to analyze the differences between moderator levels. These are provided by the *metafor* package (Viechtbauer, 2010) and the *clubSandwich* package (Pustejovsky & Pustejovsky, 2020). The mean effect size was computed for every level of each moderator variable. The categorical moderators were dummy-coded and entered as predictors in regression models. To avoid multicollinearity, we investigated each moderator in a separate analysis unless stated otherwise. There were several moderator levels with effect sizes from less than five studies per level. These moderators were not further inspected because the results could not be trustworthy.

We included continuous as well as categorical moderators in our analyses. Continuous moderators were square-root-transformed to obtain distributions closer to normal distributions (Fidell & Tabachnick, 2003, p. 135). The level with the fewest observations was used as the reference level for the categorical moderators. The levels of the categorical moderators, inserted as the predictors in the regressions, indicated whether they significantly differed from the reference level. We computed the overall significance of the moderator by using the Welch-Test of Moderators and tested for the remaining residual heterogeneity with the Q-test.

We coded 27 instructional methods. Compared to the total number of effect sizes, most methods provided only a few effect sizes and occurred in different combinations. So, analyzing each instructional method as a moderator would have produced less reliable results and enabled less clear interpretations. We first conducted a latent class analysis to test for influences of the instructional methods. With this analysis, subgroups within the database could be identified. We then tested the effectiveness of interventions in each latent class of instructional methods. Additionally, we examined our theoretical matrix in which the instructional methods were assigned. The effects of the instructional methods at each level were aggregated to determine potential differences.

7.4 Results

7.4.1 Publication Bias

Figure 29 shows the funnel plot of the included effect sizes. A publication bias in the data would indicate an underrepresentation of effect sizes close to zero. No evidence for a publication bias was found in the funnel plot. However, the result of the Egger regressions indicated such an underrepresentation of small effect sizes. This should be kept in mind when interpreting the results of this meta-analysis.

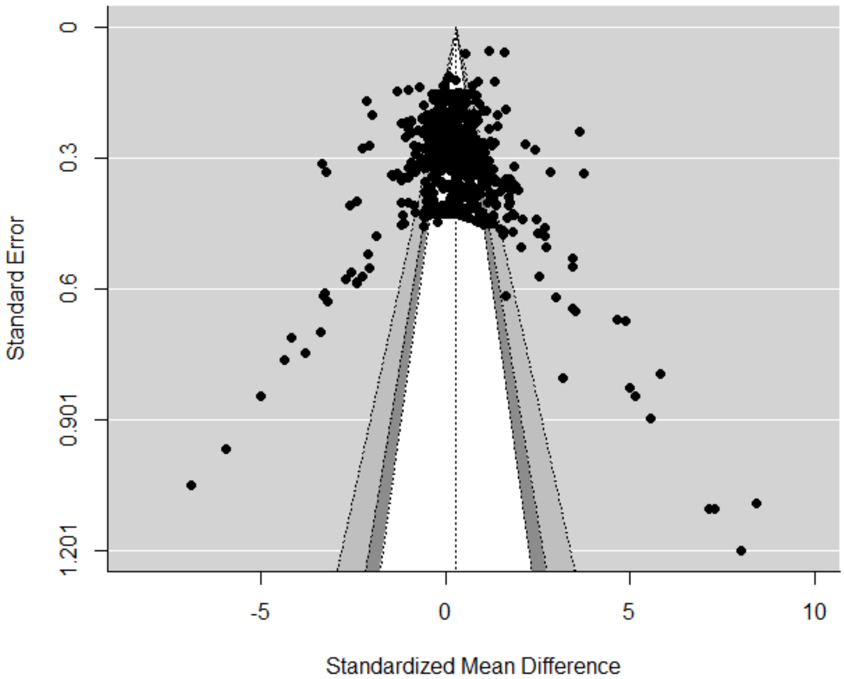
7.4.2 Study Characteristics

Findings from 151 studies nested in 138 articles with 797 effect sizes in total were meta-analytically aggregated (overall N = 16589; for the list of included studies, see Appendix H). The median sample size for the studies was 43, and the sample mean age ranged from 4.17 to

17.45 years. There were 13 studies with children in kindergarten or preschool. Children in primary school participated in 51 studies. Additionally, we included 80 studies from secondary school and four with participants from several school levels. Seventy-five studies were conducted in North America, followed by 40 from Europe, 25 from Asia, and eight from Australia. Two studies came from South America and one from Africa.

Figure 29

Funnel Plot of the 797 Included Effect Sizes (Study 3)



Note. Funnel Plot of included effect sizes. The standard error on the y-axis and the standardized mean difference (Hedges' *g*) on the x-axis. White-colored area – 90% pseudo confidence interval region; dark grey colored – 95% pseudo confidence interval region; light gray colored area – 99% pseudo confidence interval region.

Seventy studies carried out interventions in the knowledge domain of mathematics. Twenty, 17, and 10 had been conducted in physics, biology, and first language, respectively. Other domains were sports (seven studies), geosciences (six studies), chemistry (four studies), engineering (four studies), and foreign language (three studies). Six and two studies could be assigned to the levels other and several.

The mean number of sessions in a transfer intervention was 6.98, and their mean duration was 47.9 minutes per session. Between the end of the transfer intervention and the transfer test was a mean time lag of 3.333 days, ranging from zero to 16.7. Thirty-eight studies provided information on interventions that were carried out for single participants. Ninety-six interventions included groups, and four studies several group sizes.

In 27 studies, a theoretical foundation was explicitly stated. Situated Learning Theory was mentioned five times. Abstraction, Identical Elements, and Analogical Transfer followed closely with four mentions each. Formal Discipline and Preparation for Future Learning each occurred in one study. Six studies were based on several theories.

To assess whether there is a risk of bias present in the primary studies, we inspected each study based on several questions based on RoB2 (Sterne et al., 2019). The results of the Risk of Bias assessment are displayed in Tables 25 and 26 in Appendix K.

7.4.3 Effect Size Characteristics

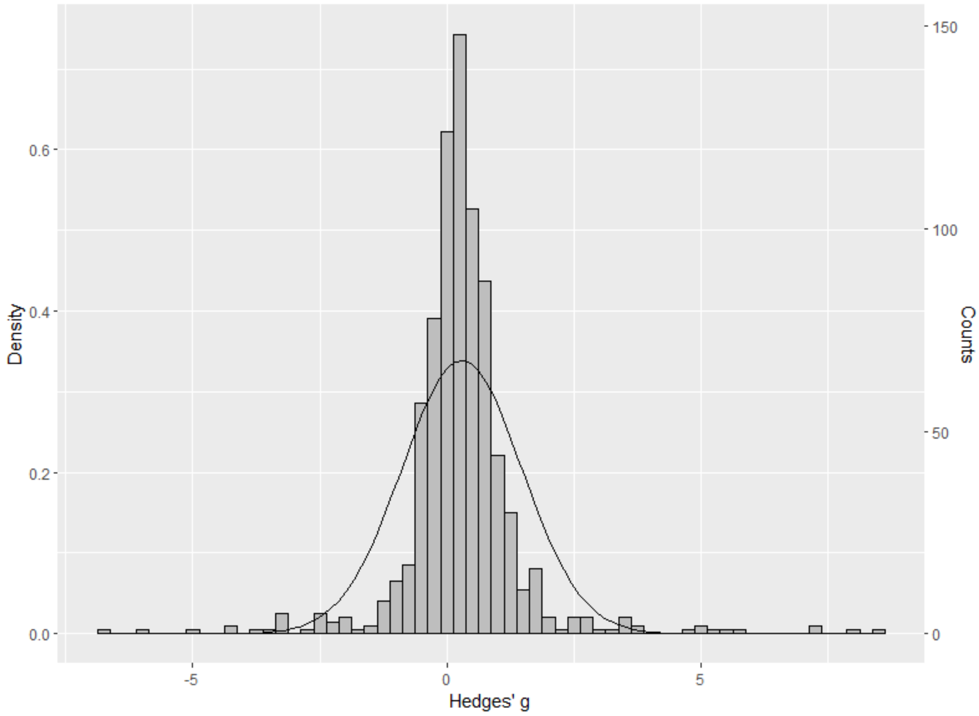
We examined the included 797 effect sizes on their distribution compared to the normal distribution (see Figure 30). These effect sizes ranged from g = -6.874 to g = 8.409, with a median of g = 0.239 and a mean of g = 0.292. The skewness of the distribution was 0.927. This means that the tail on the right side of the distribution extends towards more positive values. Therefore, the distribution was not perfectly symmetrical. The kurtosis was 16.369, which means that the distribution was leptokurtic. Therefore, it is more likely to produce outliers than a distribution with a smaller kurtosis. In conclusion, the distribution of the effect sizes did not fit perfectly into the normal distribution.

7.4.4 Main Meta-Analytic Results

7.4.4.1 Knowledge Transfer After Interventions. As expected, the meta-analytic mean Hedges' g was positive. It had a value of $g^+ = 0.285$ ($p < .001$, 95% CI [.175; .395]), which indicated that students were 0.285 standard deviations more successful in transfer measures after transfer interventions than after control interventions. However, the effect found is smaller than expected. This finding confirmed the main hypothesis H1 that instructional transfer interventions are more effective in promoting knowledge transfer than control interventions averaged over studies and instructional methods with the restriction of a small instead of a medium effect found (Cohen, 1988).

Figure 30

Distribution of the Included Effect Sizes (Study 3)



Note. Distribution of included effect sizes (k = 797) in comparison to the normal distribution. Density – Density of distribution, Counts – Number of effect sizes.

7.4.4.2 Heterogeneity. The heterogeneity in this meta-analysis comprised two variance components. These were the within-study and the between-study variance (Assink & Wibbelink, 2016). In our model, the within-study variance was σ^1 = .386, and the between-study variance σ^2 = .301. Approximately 8.95% of the total variance could be attributed to sampling variance, whereas 51.15% was attributed to differences between effect sizes within studies (within-study variance). The remaining 39.90% of the total variance resulted from differences between studies (between-study variance). According to Hunter and Schmidt (1990), if the proportion of sampling variance comprises less than 75%, there are potential moderating effects because of the substantial variation between effect sizes (Assink & Wibbelink, 2016). The significant *Q*-test for heterogeneity (Q = 6639.022, df = 796, p < .001) and the high I^2 index (I^2 = 91.05) also indicated a substantial heterogeneity of the effect sizes. This might be due to the influences of moderator variables (Higgins et al., 2003).

7.4.5 Moderator Analyses

In the following, we describe the results of the moderator analyses. We inspected each moderator in a separate analysis unless stated otherwise. Details on subgroup and moderator analyses for each moderator level can be found in Table 22 in Appendix J.

7.4.5.1 Instructional Methods. There were enough effect sizes for 20 of the initial 27 instructional methods in the primary studies to conduct moderator analyses. This means that at least five studies provided information on that instructional method. So, the moderator analyses did not further inspect the methods of *Advance Organizer*, *Generate Drawings, Interleaved Practice*, *Mnemonic Systems*, *Paraphrasing*, *Problem Posing*, and *Rehearsal*.

The instructional methods differed descriptively in their effectiveness on knowledge transfer. The effect sizes for the instructional methods as subgroups ranged from g^+ = 0.032 for *Verbal or Written Instructions* (95% CI [-.602; .665]) to g^+ = 0.567 for *Examples of Concepts* (95% CI [.088; 1.046]). We also inspected the instructional methods simultaneously in a meta-regression. All instructional methods were equally effective compared to the reference level *Verbal or Written Instructions*. Additionally, the number of methods had no significant effect on knowledge transfer (g^+ = 0.055, p = .063).

7.4.5.2 Cognitive Engagement, Abstraction, and Concretization. We tested whether the predictions of the modes of cognitive engagement from the ICAP framework applied to the mean effect of instructional methods on knowledge transfer. Table 14 provides an overview of these mean effects in the matrix. In total, 69 studies were categorized as containing only passive methods, whereas 40, 17, and seven studies were categorized as solely including active, constructive, and interactive methods, respectively. Moreover, we allocated 48 studies with methods that included exclusively the cognitive process of *Abstraction* and 57 with only *Concretization* to the matrix. The ICAP framework and our second hypothesis H2 predicted that interactive instructional methods should be most effective, followed by constructive, active, and passive methods. However, the analysis indicated no significant differences in the effects between the four levels. Descriptively, active methods showed the highest effectiveness with Hedges' g^{+}_{MEAN} of 0.305 (95% CI [.086; .524]), followed by interactive methods (g^{+}_{MEAN} = 0.289, 95% CI [-.040; .617]) and constructive methods (g^{+}_{MEAN} = 0.285, 95% CI [.126; .445]). Passive methods were the least effective, with a g^{+}_{MEAN} of 0.269 (95% CI [.140; .398]). Hypothesis H2 was not confirmed.

We also examined the third hypothesis H3; instructional methods that foster *Abstraction* of knowledge promote transfer better than methods that foster *Concretization*. This hypothesis must be rejected as well. Instructional methods that foster *Abstraction* (g^{+}_{MEAN} = 0.288, 95% CI [.159; .417]) were similarly effective as instructional methods that foster *Concretization* (g^{+}_{MEAN} = 0.306, 95% CI [.165; .447]).

Last, we tested whether the combination of concrete passive and active methods as well as the combination of abstract constructive and interactive methods were more effective in fostering transfer than their counterparts (Hypothesis H4). For this, we combined the mean effectiveness of the matrix into four categories: First, concrete passive and active instructional methods; second, concrete constructive and interactive methods; third, abstract passive and active methods; and fourth, abstract constructive and interactive methods. Abstract constructive and interactive instructional methods were less effective (g^{+}_{MEAN} = 0.261, 95% CI [.125; .340]) than abstract passive and active methods (g^{+}_{MEAN} = 0.299, 95% CI [.128; .469]). The other comparison was impossible because there were no methods with enough

Table 14

Mean Effect Sizes of Instructional Methods in Theoretical Matrix

	Abstraction						Concretization						Total g^+_{MEAN}					
	Method	j	k	g^+	p	95% CI	Method	j	k	g^+	p	95% CI		j	k	g^+	p	95% CI
Passive	Advance Organizer	2	13	-	-	-	Annotated Examples	28	76	0.368	***	[.163; .572]						
	Explanation	25	120	0.232	*	[.050; .415]	Concrete Models	11	51	0.556	*	[.003; 1.109]						
	Multiple Representations	22	186	0.543	***	[.262; .823]	Demonstration	21	89	0.456	*	[.401; .871]						
	Scaffolding	14	63	0.148	*ns*	[-.078; .373]	Examples of Concepts	19	100	0.567	*	[.088; 1.046]	**Passive**	104	505	0.269	***	[.140; .398]
							Feedback	18	81	0.116	*ns*	[-.302; .533]						
							Verbal or Written Instructions	10	63	0.032	*ns*	[-.602; .665]						
							Visualizations	16	67	0.347	*ns*	[-.078; .772]						
Active	Concreteness Fading	7	33	0.103	*ns*	[-.408; .614]	Drill	6	59	0.235	*ns*	[-.377; .846]						
	Interleaved Practice	1	4	-	-	-	Imitation	5	28	0.132	*ns*	[-.134; .398]						
	Variability of Practice	25	153	0.314	*ns*	[-.050; .678]	Mnemonic Systems	1	4	-	-	-						
							Movement	10	34	0.378	**	[.125; .630}	**Active**	59	295	0.305	**	[.086; .524]
							Paraphrasing	1	1	-	-	-						
							Realistic Problems	9	30	0.666	*	[.134; 1.198]						
							Rehearsal	2	6	-	-	-						
Constructive	Comparison of Annotated Examples	9	27	0.160	*ns*	[-.232; .551]	Generate Drawings	4	32	-	-	-						
	Comparison of Concepts	8	55	0.416	*ns*	[-.014; .845]	Problem Posing	4	10	-	-	-	**Construc-tive**	38	169	0.285	***	[.126; .445]
	Self-Explanations	21	94	0.289	**	[.082; .495]												
Inter-active	Collaborative Learning	10	32	0.289	*ns*	[-.040; .617]							**Inter-active**	10	32	0.288	*ns*	[-.040; .617]
Total g^+_{MEAN}	**Abstraction**	116	630	0.288	***	[.159; .417]	**Concretization**	102	459	.306	***	[.165; .447]						

Note. Matrix of ICAP framework's modes of cognitive engagement (Passive, Active, Constructive, and Interactive) and Abstraction, as well as Concretization. 27 instructional methods are assigned to categories based on their regular use in interventions; g^+ – Hedges' g of instructional method; p – p-value (* = $p < .05$, ** = $p < .01$, *** $p < .001$); 95% CI – 95% Confidence interval; g^+_{MEAN} – mean of Hedges' g of the level. Missing values are indicated by -. This means that the respective analysis was not possible because the number of effect sizes on the study level for this analysis was too small.

effect sizes to compute the group of concrete constructive and instructional methods. Hypothesis H4 must be rejected.

7.4.5.3 Latent Class Analysis. Based on the investigation of the instructional methods, two conclusions can be drawn beforehand: First, several instructional methods were not found often enough in studies to conduct moderator analyses. Second, none of the included instructional methods differed from the reference level *Verbal or Written Instructions* in the total moderator analysis.

Most of the instructional methods were used in combination with at least one other instructional method in the interventions. Some instructional methods were combined particularly often: *Explanations* and *Multiple Representations* were each found in combination with 21 other instructional methods. *Examples of Concepts*, *Demonstration,* and *Variability of Practice* occurred with 19 other methods. In 56 effect sizes, *Demonstration* and *Explanations* were used together in interventions. Following closely, *Explanations* and *Examples of Concepts* were combined in 55 effect sizes. The collective use of *Multiple Representations* and *Visualizations* was found in 30 effect sizes.

Latent classes usually determine subgroups based on participants' characteristics. Instead, we used latent classes to form subgroups of included studies based on the applied instructional methods coded dichotomously as *yes* or *no*. Each latent class contained multiple studies with several instructional methods. We then compared the classes as moderator levels.

The estimation of the model was conducted several times, each with a different number of specified classes. The number of classes ranged from one to 10 for this estimation. For each model, the best log-likelihood value was replicated multiple times with the initial specification of 30 random starts and a maximum of 3000 iterations. Table 15 presents the model fit indices for the estimated models with one to five classes; the fit indices for the models with more than five classes indicated a worse fit. The models with two, three, and four classes had entropies above .80, which is considered high (Clark & Muthén, 2009).

The BIC (Bayesian Information Criterion), regarded as the most reliable indicator of model fit (Nylund et al., 2007; Vermunt, 2002; Weller et al., 2020), was lowest for the model with four classes, whereas the sample-size adjusted BIC was lowest for the model with five classes. The consistent AIC (Akaike Information Criterion) was lowest for the model with three classes. The entropy was highest for the model with four classes with .915. Thus, we chose the model with four classes. It fitted the data approximately well.

Table 15

Fit Indices for The Latent Transitions Model With One to Five Latent Classes

	One class	Two classes	Three classes	Four classes	Five classes
Loglikelihood	-4904	-4641	-4487	-4391	-4324
resid. df	771	744	717	690	663
BIC	9982	9637	9509	9497	9581
Sample-size adjusted BIC	9900	9469	9255	9158	9155
cAIC	10008	9690	9589	9604	9715
Entropy	-	.804	.851	.915	NaN

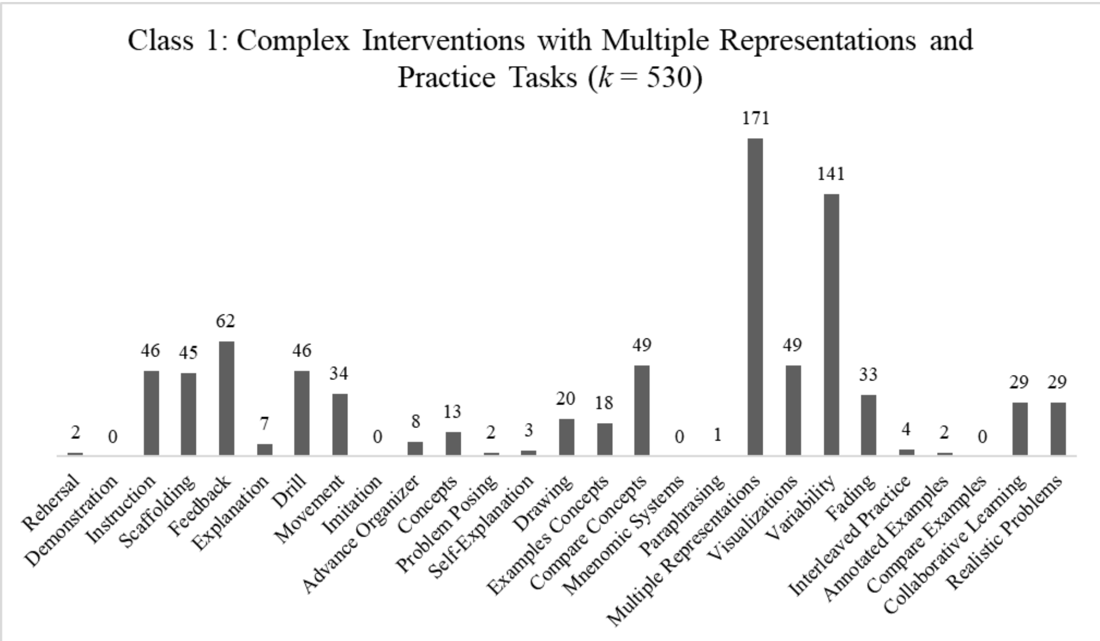
Note. resid. df – residual degrees of freedom; BIC – Bayesian Information Criterion; cAIC – consistent Akaike Information Criterion.

The first latent class included 530 effect sizes from 99 studies (e.g., Mullins et al., 2011). Effect sizes in this class were from transfer interventions with a broad range of instructional methods (see Figure 31). Most interventions used *Multiple Representations* (k = 171) and *Variability of Practice* (k = 141) as instructional methods. Therefore, this class was labeled *Complex Interventions with Multiple Representations and Practice Tasks*. The mean effect of transfer interventions compared to control interventions in this class was g^+ = 0.214 (p = .001, 95% CI [.084; .343]). The second class contained effect sizes from transfer interventions with diverse instructional methods and predominantly *Demonstration* (k = 77) and *Explanations* (k = 58) (see Figure 32, e.g., Katzlberger, 2005). Eighty-six effect sizes from 25 studies were assigned to this class, *Complex Interventions with Demonstrations and*

Explanations. It yielded a mean effect of $g^+ = 0.448$ ($p = .016$, 95% CI [.083; .814]). Effect sizes from transfer interventions that mostly relied on the method *Self-Explanations* were assigned to the third class, *Self-Explanation* (see Figure 33). This resulted in 115 effect sizes from 31 studies (e.g., Rittle-Johnson, 2006). Here, the mean effect size was $g^+ = 0.315$ ($p <$ .001, 95% CI [.157; .472]). Last, the fourth class included 66 effect sizes from transfer interventions that mainly relied on *Examples of Concepts* (k = 59) and *Explanations* (k = 53), resulting in the latent class *Explanations and Examples of Concepts* (see Figure 34, e.g., Kyun & Lee, 2009). The mean effect for these 66 effect sizes was $g^+ = 0.315$ ($p < .001$, 95% CI [.157; .472]).

Figure 31

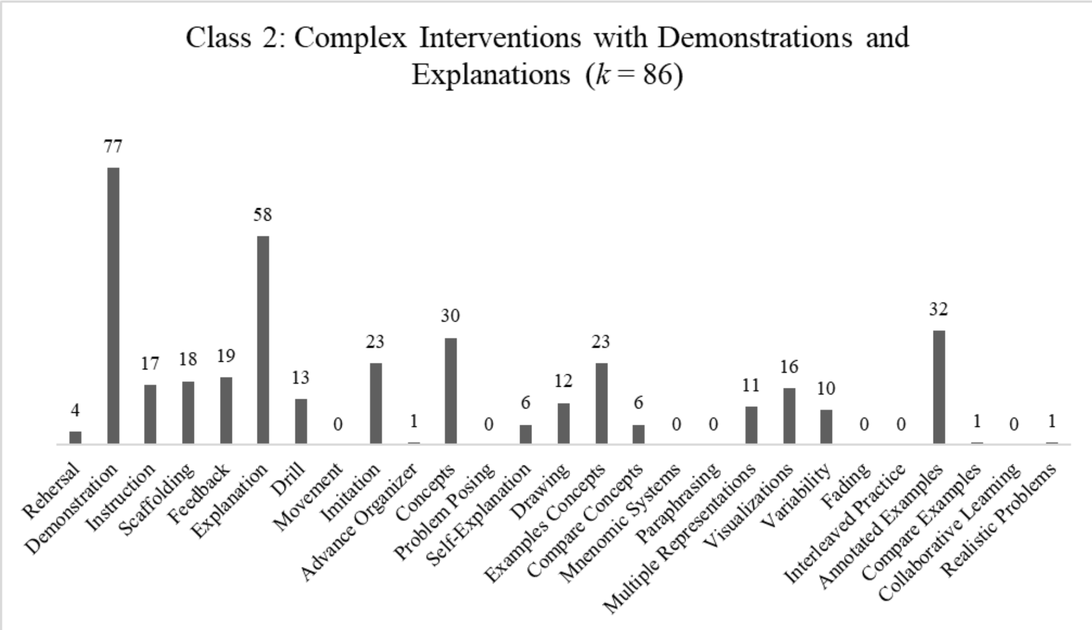
Overview Latent Class 1 (Complex Interventions With Multiple Representations and Practice Tasks)



Note. Overview of class 1 (*Complex Interventions with Multiple Representations and Practice Tasks*) of the four latent classes. The height of the bars and the number above each bar represents the number of effect sizes in this class for the respective instructional methods.

Figure 32

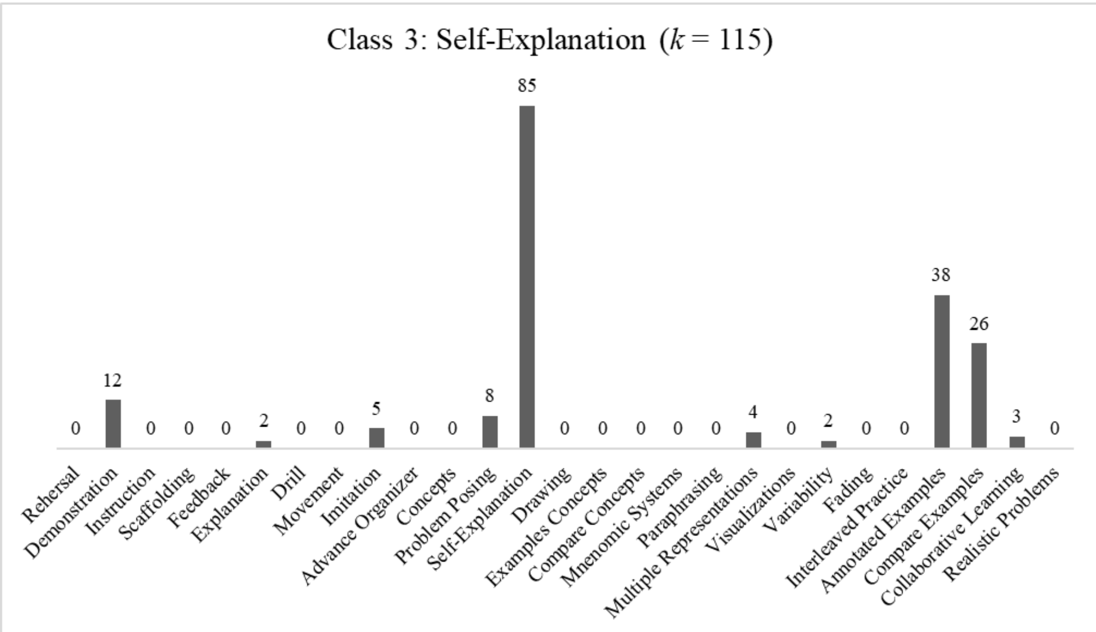
Overview Latent Class 2 (Complex Interventions With Demonstrations and Explanations)



Note. Overview of class 2 (*Complex Interventions with Demonstrations and Explanations*) of the four latent classes. The height of the bars and the number above each bar represents the number of effect sizes in this class for the respective instructional methods.

Figure 33

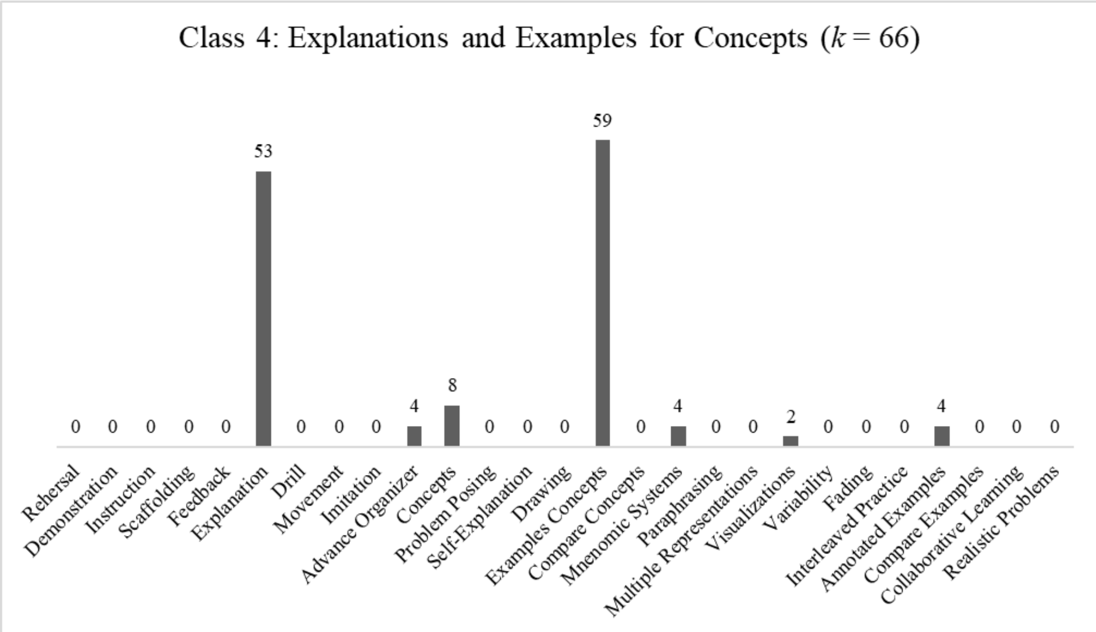
Overview Latent Class 3 (Self-Explanation)



Note. Overview of class 3 (*Self-Explanation*) of the four latent classes. The height of the bars and the number above each bar represents the number of effect sizes in this class for the respective instructional methods.

Figure 34

Overview Latent Class 4 (Explanations and Examples of Concepts)



Note. Overview of class 4 (*Explanations and Examples of Concepts*) of the four latent classes. The height of the bars and the number above each bar represents the number of effect sizes in this class for the respective instructional methods.

The determination of latent classes is based on the probability of the estimated effect sizes' class membership. This percentage describes the probability of an effect size being assigned to one of the four classes. The estimated class membership was 66.50% in *Complex Interventions with Multiple Representations and Practice Tasks*. This means that an effect size had a probability of 66.50% to be assigned to the first class. The membership was 10.79% in *Complex Interventions with Demonstrations and Explanations*, 14.43% in *Self-Explanation,* and 8.28% in *Explanations and Examples of Concepts.* With this distribution of cases, the latent class analysis complied with two general requirements (Weller et al., 2020): Each latent class of the analysis included more than 50 effect sizes (Muthén & Muthén, 2000) and more than 5% of the sample (Shanahan et al., 2013).

The four latent classes did not differ in their results regarding the effectiveness of interventions. Compared to the class *Explanations and Examples of Concepts* ($g^{+}_{MOD} = 0.233$)

as the reference level, the other classes were descriptively more effective but statistically nonsignificant. *Complex Interventions with Multiple Representations and Practice Tasks* yielded slightly larger effects ($g^+_{MOD}=0.260$). *Complex Interventions with Demonstrations and Explanations* ($g^+_{MOD} = 0.373$) and *Self-Explanation* ($g^+_{MOD} = 0.327$) showed even more substantial, but not statistically significant, effects. Information on effect sizes and covariates of the four latent classes is presented in Table 16.

Table 16

Mean Effect Sizes and Covariates for the Four Latent Classes

	j	*k*	g^+	*p*	95% CI	Num.	Dur.	Time betw.	Total *N*	Age exp.	Age comp.
Class 1: *Complex Interventions with Multiple Representations and Practice Tasks*	99	530	0.214	**	[.084; .343]	7.59	47.79	9.92	67.40	10.64	10.64
Class 2: *Complex Interventions with Demonstrations and Explanations*	25	86	0.448	*	[.083; .814]	12.74	44.03	9.22	60.45	9.93	10.03
Class 3: *Self-Explanation*	31	115	0.315	***	[.156; .472]	5.00	42.00	6.01	56.64	13.05	13.06
Class 4: *Explanations and Examples of Concepts*	12	66	0.258	*ns*	[-.014; .529]	3.65	39.13	2.14	62.52	14.98	14.97

Note. Overview of latent classes; *j* – Number of studies; *k* – Number of effect sizes; *g* – mean Hedges' *g*; *p* – *p*-value (* = $p < .05$, ** = $p < .01$, *** $p < .001$); 95% CI – Confidence Interval; Num. – Mean number of sessions in the level; Dur. – Mean duration of intervention sessions in minutes; Time betw. – Number of days between last session and transfer assessment; Total *N* – Mean total N of participants; Age exp. – Mean age of participants in experimental group; Age comp. – Mean age of participants in comparison group.

7.4.5.1 Near and Far Transfer. We examined to what extent the transfer distance moderated the effectiveness of transfer interventions compared to control interventions. For this, we used the Taxonomy for Far Transfer (Barnett & Ceci, 2002). In this moderator analysis, we analyzed the influence of the transfer distance operationalized as six context factors:

Knowledge Domain, *Physical Context*, *Temporal Context*, *Functional Context*, *Social Context,* and *Modality*. Most of the included studies assessed only near transfer. The context factor with the most assessments on far transfer was *Modality* (k = 290), followed by *Temporal Context* (k = 146), *Knowledge Domain* (k = 44), and *Social Context* (k = 43). *Physical Context* and *Functional Context* were assessed the least as far transfer with k = 13 and k = 7 from four and two studies, respectively. Therefore, the following moderator analyses did not inspect these context factors.

There were no statistically significant differences between near and far transfer assessments as moderator levels in the effectiveness of interventions. The mean effect of near transfer assessments ranged from $g^{+}_{MOD} = 0.221$ for *Modality* to $g^{+}_{MOD} = 0.283$ for *Knowledge Domain.* Assessments of far transfer yielded slightly stronger but nonsignificant effects in general compared to the near transfer assessments ($g^{+}_{MOD} = 0.261$ for *Knowledge Domain* to $g^{+}_{MOD} = 0.419$ for *Social Context*). Therefore, hypothesis H5a must be rejected.

We also investigated the differences between the six context factors concerning the effectiveness of the interventions. We included the four context factors (*Knowledge Domain*, *Temporal Context*, *Social Context,* and *Modality*) as predictors in a meta-regression. This moderator analysis did not yield significant results (Test of Moderators, F ($df1$ = 4, $df2$ = 473) = 1.073, p = .369). There were no differences between the dimensions and hypothesis H5b, the effectiveness of interventions differs along the six transfer dimensions, must be rejected.

7.4.6 Exploratory Moderators

In the following, we present the results of the exploratory research questions. First, we analyzed the interactions between the four latent classes of instructional methods and the six context factors of the transfer distance. Furthermore, the other exploratory moderators that we inspected are described in short. These are sample characteristics, intervention characteristics as well as characteristics of the method.

7.4.6.1 Interaction Between Latent Classes and Transfer Distance. The four latent classes did not differ significantly in their effect on near knowledge transfer. This applied to all six context factors (see Figure 48 in Appendix I). However, there were descriptive differences between the latent classes in the assessments of far knowledge transfer: For the context factor *Knowledge Domain*, the class *Complex Interventions with Multiple Representations and Practice Tasks* yielded negative results on far transfer. For this class, transfer interventions promoted far transfer less than control interventions. In the other three classes, transfer interventions were more effective than control interventions, and there were no or only minor differences between near and far transfer. The class *Explanations and Examples of Concepts* discerned from the other classes for the context factor *Temporal Context*, providing more positive results on far than near transfer assessments. The remaining three classes promoted near transfer slightly more than far transfer. All transfer interventions were more effective than control interventions. For the context factor *Social Context*, the class *Complex Interventions with Demonstrations and Explanations* yielded negative results for far transfer compared to control interventions. Furthermore, the class *Explanations and Examples of Concepts* promoted far transfer slightly more than near transfer. The other two classes fostered near and far transfer similarly more than control interventions. On the context factor *Modality*, the class *Complex Interventions with Demonstrations and Explanations* was more effective for far than near transfer. Conversely, for the class *Explanations and Examples of Concepts*, transfer interventions were less effective than control interventions for far transfer. The other two classes did not show differences in their promotion of near and far transfer. Both yielded positive results for the two transfer distances. *Physical Context* and *Functional Context* were comparable for all latent classes because of too few effect sizes for far transfer ($k = 13$ and $k = 7$). In these, all latent classes promoted far transfer less than near transfer; but they were more effective than control interventions.

7.4.6.2 Other Exploratory Moderators. We investigated several characteristics of the sample, the intervention, and the methodology. The mean age of the experimental group showed no influence on the effectiveness of interventions. For the comparison group, the mean age yielded significant influences. The higher the mean age of the comparison groups, the smaller the mean advantage of transfer interventions compared to control interventions ($g^{+}_{CHANGE} = -0.168$, $p = .039$, 95% CI [-.327; -.009]). The effects also differed significantly between the age groups of the experimental and control groups. Compared to children younger than six ($g^{+}_{MOD} = 0.635$, for experimental; $g^{+}_{MOD} = 0.635$, for control), the group of adolescents between 12 and 18 years showed weaker effects on knowledge transfer ($g^{+}_{MOD} = 0.152$, $p = .047$, for experimental; $g^{+}_{MOD} = 0.149$, $p = .046$, for control).

For the experimental group, the school level primary school ($g^{+}_{MOD} = 0.152$, $p < .001$) and secondary school ($g^{+}_{MOD} = 0.154$, $p < .001$) differed significantly from the level kindergarten/preschool ($g^{+}_{MOD} = 0.919$). The advantage of transfer interventions on knowledge transfer over control interventions was larger for children in kindergarten or preschool than for school students in the experimental conditions. Comparable results are found for the control groups: The mean effect for kindergarten/preschool ($g^{+}_{MOD} = 0.944$) was significantly more potent than for primary ($g^{+}_{MOD} = 0.256$, $p < .001$) and secondary school ($g^{+}_{MOD} = 0.202$, $p < .001$). Transfer interventions were significantly more effective than control interventions in cases where the comparison group contained children in kindergarten or preschool. Consistent with this result, the mean advantage of transfer interventions over control interventions was smaller the higher the class level of the experimental and the comparison group was ($g^{+}_{CHANGE} = -0.134$, $p = .047$, for experimental group; $g^{+}_{CHANGE} = -0.215$, $p = .002$, for comparison group).

As a characteristic of the intervention, the setting significantly influenced the interventions' effectiveness. Interventions in the field ($g^{+}_{MOD} = 0.311$) were more effective in fostering knowledge transfer than in laboratories ($g^{+}_{MOD} = -0.149$, p = .034). It must be stated here that the moderator level of interventions in laboratories contained only 27 effect sizes from seven studies, which limits the interpretation of the result. None of the other intervention characteristics yielded significant results.

For the methodological characteristics, only the implementation check showed significant differences between the moderator levels. For cases in which a successful implementation check was reported (g^{+}_{MOD} = 0.772), Hedges' g was significantly stronger than in cases in which the implementation check was not successful (g^{+}_{MOD} = 0.012, p < .001). The same applied to studies that did not include an implementation check (g^{+}_{MOD} = 0.318, p < .001). Transfer interventions that successfully conveyed the intended knowledge were also more successful in promoting knowledge transfer than control interventions.

7.5 Discussion

7.5.1 Main Findings

This meta-analysis investigated the effect of transfer interventions for school students on their knowledge transfer compared to control interventions. As expected in the first hypothesis H1, transfer interventions were more effective in promoting knowledge transfer than control interventions with a Hedges g^{+} of 0.285. This means that instructional methods in transfer interventions promoted knowledge transfer better than control interventions that did not include these methods. According to this, the experimental groups had a higher mean score in the assessment of knowledge transfer after interventions than the comparison groups.

The main result is in line with previous findings on knowledge transfer, for example, in work contexts (e.g., Blume et al., 2009; Cheng & Hampson, 2008; Ford & Weissbein, 1997) or studies on specific knowledge domains (e.g., Kaminske et al., 2020; Rittle-Johnson et al., 2017; Scherer et al., 2019). Kaminske and colleagues (2020) purported the importance of several instructional methods (e.g., multiple examples and interleaving) to foster knowledge transfer in biology. Also, the method of *Self-Explanations* in mathematics yielded an effect of g = 0.46 (k = 9, 95% CI [.16; .76]) for immediate and g = 0.32 (k = 6, 95% CI [.02; .63]) for delayed transfer tests in a meta-analysis (Rittle-Johnson et al., 2017). Similar results are found for the impact of programming trainings on programming as near transfer (g = 0.75, 95% CI [0.39, 1.11]) and on mathematical skills as far transfer (g = 0.57, 95% CI [0.34, 0.80]) in school students (Scherer et al., 2019). Compared to these examples of prior research on knowledge transfer, our main result is similar but smaller. This was expected because of the broader range of included studies in several knowledge domains and with several instructional methods. Moreover, we included not only true control groups but also treated groups that learned with fewer instructional methods in an intervention.

Concluding the first hypothesis, transfer interventions positively impacted school students' knowledge transfer more than groups that did not learn with these methods. Our results are comparable to prior findings, underlining the trustworthiness. The finding speaks for the importance of additionally promoting knowledge transfer in school students by providing transfer interventions in several domains.

7.5.2 Moderator Variables

The moderator analyses showed that the meta-analytical results were stable for most of the inspected moderators. Transfer interventions were more effective in promoting knowledge transfer than control interventions across instructional methods and transfer distance. Only some of the exploratory moderators significantly influenced the effect of transfer interventions.

We found no differences in the effectiveness of instructional methods based on the learners' cognitive engagement and the cognitive processes of *Abstraction* and *Concretization*. The passive, active, constructive, and interactive methods did not differ in their promotion of knowledge transfer. The second hypothesis H2 must be rejected. Contrary to our third hypothesis H3, instructional methods that foster *Abstraction* (g^+ = 0.288) were slightly less effective than those that foster *Concretization* (g^+ = 0.306). Unlike our hypothesis, there were no advantages for methods that fostered the *Abstraction* of learned contents. Also, in contrast to our fourth hypothesis H4, abstract constructive and interactive instructional methods were less effective ($g^+ = 0.261$) than abstract passive and active methods ($g^+ = 0.299$). All included categories of instructional methods were equally effective in fostering knowledge transfer. Therefore, our three moderator hypotheses H2, H3, and H4 were rejected. These results imply that instructional methods from all theoretical categories positively affected knowledge transfer. Contrary to the prediction of the modes of cognitive engagement from the ICAP framework, not only the active, constructive, and interactive but all four modes promoted knowledge transfer. Our results show no advantage for the expected cognitive engagement based on the methods or the assumed cognitive transfer process. This speaks for the possibility of successfully fostering knowledge transfer in school students with various instructional methods.

There were no significant differences between near and far transfer in the effectiveness of interventions on the context factors *Knowledge Domain*, *Temporal Context*, *Social Context*, and *Modality*. We inspected the possible influences of the transfer distance based on the Taxonomy for Far Transfer (Barnett & Ceci, 2002). Interventions were equally effective on near and far transfer assessments. However, most of the primary studies included only near transfer. In conclusion, the results of far transfer must be interpreted with caution. Previous studies showed higher effects for near than far transfer (e.g., Scherer et al., 2019). This is in alignment with the prediction in the theoretical differentiation of transfer that far transfer is only scarcely found (e.g., Perkins & Salomon, 1992). However, we did not find differences in near and far transfer. This result could be attributed to the fact that most interventions included only near transfer assessments. Moreover, we could not analyze the factors *Physical Context* and *Functional Context* because of the scarcity of effect sizes for far transfer.

The influence of intervention characteristics on knowledge transfer was stable across most of the examined exploratory moderators. Only the following exploratory moderators impacted the effectiveness: For both groups, age, school level, and class were significant predictors. The higher the age group, the school level, and the class level, the lower the mean effect of transfer interventions on knowledge transfer compared to control interventions. The mean age of the comparison groups was also a significant moderator, pointing in the same direction. This speaks for the necessity of early interventions to promote knowledge transfer more effectively. Furthermore, the setting of the intervention and the implementation check showed significant differences in the effectiveness of interventions. Interventions conducted in the field led to better results in transfer assessments. Also, when a successful implementation check was included in the study, the students had a higher mean knowledge transfer score afterward.

The exploratory moderator analyses demonstrated that the characteristics of the sample and the setting must be considered when carrying out a transfer intervention. Lower school levels, classes, and ages of participants should be preferred. Transfer interventions seem to have more effect on knowledge transfer for younger students. Furthermore, conducting a transfer intervention in a realistic setting appears helpful. Interventions that included a positive implementation check successfully taught the intended knowledge; this resulted in better applications of this knowledge in transfer assessments.

Regarding the moderator analyses, most of the inspected moderators did not show influences on the effectiveness of interventions. This leads to the assumption that transfer interventions are generally more effective than control interventions in promoting knowledge transfer.

7.5.3 Limitations

There are some methodological as well as content-related limitations in the included primary studies. Methodologically, two aspects must be mentioned: The first limitation refers to the control interventions. We included control groups that only received the same content but no instructional methods in addition to regular school instructions as well as comparison groups that participated in an intervention but learned with fewer instructional methods. The effect of transfer interventions compared to less extensive interventions is expected to be lower than compared to true control groups (e.g., Scherer et al., 2019). Therefore, it is assumable that we underestimated the general advantage of transfer interventions.

Second, most of the primary studies did not describe the details of their instructional interventions. This obstructs the assessment of the entirety of the used methods. Additionally, the absence of clear definitions and descriptions of instructional methods in the primary studies limits the generalizability. We coded the instructional methods based on definitions proven in past research. However, there are no unified definitions and classifications for instructional methods. This complicated the coding of primary studies and limits the generalizability of our results because different usages of instructional methods can lead to different cognitive engagement activities and outcomes (Chi & Wylie, 2014).

Similarly, our first content-related limitation includes questions on the definition of transfer. As mentioned before, there are several definitions of knowledge transfer. This leads to different transfer assessments. Primary studies varied in the usage of the word *transfer* and its measurement. We coded assessments that fitted the definition of knowledge transfer presented above. These included measurements labeled as transfer in primary studies and measurements that carried other labels but fitted the definitions (e.g., *application item*, Kapur, 2010). This approach yields the risk of missing single transfer assessments with other labels or including assessments labeled as transfer but not genuinely meeting the definition above.

The second content-related limitation addresses the four modes of cognitive engagement. As mentioned earlier, the modes are independent of the instruction itself (Chi & Wylie, 2014). Therefore, the same instructional method can cause different cognitive engagement activities in students. We sorted the instructional methods based on the expected cognitive engagement. However, it is possible that these methods were used differently in primary studies without addressing it. This could have caused other modes of cognitive engagement than expected based on our theoretical matrix. For example, the presentation of *Examples of Concepts* is coded as passive in our matrix. However, in a primary study, this method could be combined with underlining relevant parts of the example without addressing this in the article, making it an active method. Such other usages of the instructional methods would explain the similar effects of all four modes of cognitive engagement that we found in the moderator analysis.

7.5.4 Implications

7.5.4.1 Implications for Education. The findings provide several educational implications, especially for interventions that facilitate knowledge acquisition at school. The main result confirms the advantage of transfer interventions over control interventions in the promotion of knowledge transfer. This means that transfer interventions with instructional methods are generally more helpful in fostering knowledge transfer than control interventions. The setting of interventions was a significant moderator. Transfer interventions carried out in realistic school settings were more effective than in laboratories in promoting knowledge transfer. Additionally, the number and duration of sessions and the number of instructional methods used in interventions did not significantly influence the effectiveness of interventions on knowledge transfer. Short sessions and using single methods are comparably effective to using multiple methods in longer sessions. This aligns with the practical restrictions in school education: Knowledge transfer can be achieved by using few methods in short sessions. This speaks for the realization of short interventions that include only selected methods. Such interventions can be integrated into the regular school curriculum.

Age, school level, and class were significant moderators. The higher the age, school level, and class, the smaller the effect of transfer interventions on knowledge transfer compared to control interventions. This finding underlines the importance of interventions for younger children in kindergarten or preschool. Furthermore, because of the more flexible curriculum in kindergarten and preschool, interventions that facilitate knowledge acquisition and transfer can more easily be conducted at that time.

Additionally, there were no differences between the broad and narrow knowledge domains. Interventions in STEM domains have been researched more often (k = 643) than in other broad content areas (e.g., Language, k = 49; Social Sciences, k = 4). Within the STEM domains, interventions were comparable in their effectiveness on knowledge transfer (e.g., biology, physics). Interventions to promote knowledge transfer could be applicated in a broad range of knowledge domains at school based on the results of prior research. Other domains, for example, language or social sciences, were researched less often. There is a need for more primary studies in these areas.

The matrix of the four modes of cognitive engagement and the cognitive processes of *Abstraction* and *Concretization* provides a helpful structure for practitioners. Trainers can use this structure and the instructional methods to construct trainings that aim at knowledge transfer. The instructional methods generally promote knowledge transfer in interventions for school students. On the other hand, the number of methods in trainings had no significant influence on the effectiveness. Therefore, practitioners can choose individual methods for their training that are consistent with their needs.

7.5.4.2 Implications for Future Research. This meta-analysis also yields several implications for future research in knowledge transfer. First, the matrix used in this meta-analysis should be extended by other instructional methods that have not been included. To achieve this goal, additional research in defining instructional methods is needed. By providing explicit definitions of instructional methods, future studies can be more precise in their results. This will make it easier to aggregate results across primary studies.

Further investigations on single instructional methods compared to others can provide insights into the relative importance of instructional methods for knowledge transfer. A more profound basis can be built by systematically comparing methods with each other. This might then be a solid starting point for broader guidance for practitioners as well as further studies in the research domain of knowledge transfer.

We used a latent class analysis to cluster effect sizes based on the interventions' instructional methods. These latent classes were then used as moderator variables. This approach yielded the advantage that instead of 27 separate moderators, we only investigated four classes with enough effect sizes to report more reliable results. Other meta-analyses that include a broad range of moderators should further inspect this methodological approach.

This meta-analysis inspected interventions for school students, but it might also be interesting to sum up the evidence in adult education, apart from the work context. Studies in higher or later education might also yield varying effects of interventions on knowledge transfer. The age, school level, and class of the groups significantly influenced the effectiveness of interventions. Each of these moderators pointed in the direction that interventions were more effective for younger students. Comparing this result to adult participants could provide further results on the best timing for interventions to foster knowledge transfer across the lifespan.

7.5.5 Conclusion

In this meta-analysis, we compared experimental groups that received instructional methods in interventions to comparison groups that learned with fewer methods in their knowledge transfer afterward. Altogether, transfer interventions showed an advantage in promoting knowledge transfer in school students (g^+ = 0.285) compared to control interventions. These effects were stable across most moderator variables, such as the transfer distance. Additionally, the learners' cognitive engagement and transfer processes did not yield significant differences regarding knowledge transfer. However, when conducting transfer interventions, the age and school level of the participants, as well as the setting and the successful learning of contents, should be kept in mind. These findings demonstrate the general advantage of transfer interventions in fostering knowledge transfer in school students compared to control interventions.

8 General Discussion

This dissertation aimed to investigate three key processes associated with knowledge. These are the relations between knowledge types, motivational mediators between prior knowledge and knowledge after learning, and influences on knowledge transfer. More precisely, we investigated the following questions: How are conceptual and procedural knowledge longitudinally related in mathematics? Do motivational variables mediate between prior knowledge and knowledge after learning? How strongly do instructional methods foster knowledge transfer in interventions for school students? Three meta-analyses were conducted to answer these questions. Six key insights are drawn from the results of the studies in this dissertation. The discussion starts with presenting the key insights (Chapter 8.1). It is followed by practical implications for education (Chapter 8.2) and recommendations for future research (Chapter 8.3). Last, a short take-home message in the form of a conclusion complements the content of this general discussion (Chapter 8.4).

8.1 Key Insights

The results of the three studies can be broadly summarized in six general key insights. The first key insight consists of the finding that there are significant cross-sectional as well as longitudinal relations between conceptual and procedural knowledge in the domain of mathematics. Second, prior conceptual as well as procedural knowledge in mathematics predicts later knowledge in the same knowledge type. The third key insight shows that the mediation of the relation between prior knowledge and later knowledge by motivational variables differs based on the assessed knowledge type. A fourth key insight presents the advantages of promoting knowledge transfer in young learners. Fifth, the importance of near transfer compared to far transfer in school settings is discussed. By conducting a latent class analysis, it is possible to include classes of instructional methods as moderating variables in a meta-analysis, providing more profound results than individually analyzing each method. This process is included as the sixth key insight of this dissertation.

In the following, each of the six key insights are briefly explained. They build on the results of the three studies presented in this dissertation and yield further general outcomes. The first two key insights are based on Study 1, the third is drawn from Study 2, and the last three are results from Study 3.

8.1.1 Key Insight 1: Relations Between Conceptual and Procedural Knowledge in Mathematics

Study 1 demonstrated that conceptual and procedural knowledge in mathematics were connected via bidirectional predictive relations. These relations were strong (Orth et al., 2022) and similarly high (β_{C1P2} = .270, 95% CI [.221, .318]; β_{P1C2} = .220, 95% CI [.172, .268]). Additionally, conceptual and procedural knowledge showed significant cross-sectional relations at measurement points one (r_{C1P1} = .398, 95% CI [.354; .443]) and two (r_{C2P2} = .243, 95% CI [.113; .354]). Both knowledge types were significantly interconnected cross-sectionally as well as longitudinally.

This finding contributes to the debate on whether conceptual and procedural knowledge can be measured as independent knowledge types or if they are too closely interconnected (Baroody et al., 2007; Faulkenberry, 2013; Lenz et al., 2020; Lenz & Wittmann, 2020). There is a significant cross-sectional relationship between the knowledge types. However, it is feasible to interpret the correlations as small to medium (Cohen, 1988). Based on the size of the correlations found in Study 1, the knowledge types can be measured as partially separate constructs.

This aligns with the empirical findings by Lenz and colleagues (2020), who developed an instrument that differentiated between conceptual and procedural knowledge by only using tasks attributed to one of the knowledge types. By considering cognitive processes associated with conceptual and procedural knowledge, the authors adapted the taxonomy for learning, teaching, and assessing (Anderson & Krathwohl, 2001) and specified knowledge elements decisive for each knowledge type in fraction addition and subtraction. They concluded that most previous tasks could be solved using conceptual as well as procedural knowledge and that they assessed only specific aspects of each knowledge type (Lenz et al., 2020). According to the authors, developing instruments in mathematics that successfully differentiate between conceptual and procedural knowledge is possible.

Because of the statistically significant correlations between conceptual and procedural knowledge at measurement points one and two in Study 1, advancements in measurements of the individual knowledge types are essential for further research. Improving the accuracy of measurements makes more precise interpretations of the results and the relations between conceptual and procedural knowledge possible. These measurements should consider specific tasks for the knowledge types (Gabriel et al., 2013; Lenz et al., 2020). Additionally, Faulkenberry (2013) argues that not only the selection of tasks but also the strategies employed on them are important to consider. Students can employ different strategies on the same task based on conceptual or procedural knowledge. For example, the magnitude comparison task, in which two fractions are compared, is often considered a conceptual task (e.g., Hecht & Vagi, 2012). Nonetheless, using rote strategies like comparing cross-products, the result could be interpreted as procedural knowledge (Boston et al., 2003; Faulkenberry, 2013; Faulkenberry & Pierce, 2011). Researchers must consider these differences in strategy use when constructing instruments that differentiate conceptual and procedural knowledge.

The explanatory power of this finding is currently limited to the knowledge domain of mathematics. However, more insights are possible by improving definitions and measurements of conceptual and procedural knowledge in other knowledge domains. If knowledge can be differentiated into conceptual and procedural knowledge, it will also be related cross-sectionally and longitudinally in other knowledge domains.

8.1.2 *Key Insight 2: Conceptual and Procedural Knowledge Predict Themselves in Mathematics*

Study 1 showed that conceptual and procedural knowledge predicted themselves longitudinally. We found significant relations from conceptual knowledge at measurement point one to measurement point two (β_{C1C2} = .402, 95% CI [.346; .448]) and procedural knowledge at measurement point one to measurement point two (β_{P1P2} = .336, 95% CI [.291; .380]). The finding speaks for the medium stability of both knowledge types. Prior conceptual knowledge predicted later conceptual knowledge, and prior procedural knowledge predicted later procedural knowledge. This means that children with more prior knowledge showed more later achievement in the same knowledge type.

This result is in alignment with previous findings (Bempeni & Vamvakoussi, 2014, 2015; Hallett et al., 2010; Hecht & Vagi, 2010, 2012; Lenz & Wittmann, 2020). For example, Hecht and Vagi (2010) identified four profiles of conceptual and procedural knowledge in fourth-grade students: Lower concepts as expected, lower concepts and higher procedures as expected, higher concepts and lower procedures as expected, and higher concepts and higher procedures as expected (Hecht & Vagi, 2012). These knowledge profiles showed little stability from fourth to fifth grade. By improving conceptual and procedural knowledge, children with different knowledge profiles can profit from instructions and increase their knowledge (Gabriel et al., 2013; Hallett et al., 2010, 2012).

Lenz and Wittmann (2020) speak for the necessity of differentiated considerations of knowledge profiles to derive concrete support strategies. Children with lower procedural knowledge but sufficient conceptual knowledge can rely on the latter, whereas children with low procedural and conceptual knowledge need further support (Lenz & Wittmann, 2020). Individual differences in knowledge types also adhere to differences in teaching. Therefore, teachers should assess children's prior conceptual and procedural knowledge to adjust their teaching to individual needs.

According to Study 1, conceptual and procedural knowledge predicted themselves longitudinally. This finding adds to the existing research on the influence of prior knowledge on learning outcomes (e.g., Simonsmeier et al., 2022). It is vital to foster conceptual and procedural knowledge in mathematics education because both knowledge types influence later achievement in the same knowledge type.

Limiting the generalization, the relations in Study 1 are restricted by the reliability of the measurements. Both knowledge types tend to be interconnected, making it challenging to measure each of them without the influences of the other (Baroody et al., 2007). Nonetheless, the correlations between the knowledge types found in Study 1 are low to medium size (see Chapter 8.1.1). It is possible to conclude that the measurements sufficiently captured both knowledge types.

More research is needed to answer whether this finding applies to other knowledge domains as well. The differentiation in conceptual and procedural knowledge is less researched outside of the domain of mathematics. Additionally, whether these knowledge types are separable and whether such differentiation is essential in other knowledge domains is mostly unknown.

8.1.3 Key Insight 3: Mediation of Motivational Variables on the Relation Between Prior Knowledge and Learning Depends on the Knowledge Type

Study 2 showed that the influence of motivational mediators differed based on the assessed knowledge type of the learned contents. The meta-analysis differentiated between declarative knowledge and procedural knowledge. For the mediator interest, the total effect was stronger for declarative knowledge (r_{MED} = .210, 95% CI [.194; .233]) than for procedural (r_{MED} = .030, 95% CI [-.040, .100]). The inverse was found for self-concept (r_{MED} = .086, 95% CI [.050; .127] for declarative knowledge; r_{MED} = .180, 95% CI [.147; .213] for procedural knowledge). However, only two studies provided information on interest in declarative knowledge. This limited the explanatory power of the result.

By considering the two knowledge types, interest seems to be a more crucial mediator for declarative knowledge, whereas self-concept yielded significantly more influence on procedural knowledge. Both results appear to be plausible: Interest helps in acquiring declarative knowledge, which includes facts and concepts. If learners find topics interesting, they spend more time on activities related to that field (Tobias, 1994) and acquire concepts and facts easier in that domain. Self-concept, on the other hand, benefits the execution of procedures. Learners with a strong self-concept have a positive image of themselves which predicts how the person performs procedures (Bong & Skaalvik, 2003; Rosenberg, 1979).

In examining text comprehension, several studies inspected the impact of interest on declarative knowledge. Text comprehension can be categorized as declarative knowledge if operationalized as explicit content knowledge (McCarthy & McNamara, 2021). Schiefele (1990, 1992) examined the influence of interest-related topic knowledge on text comprehension. In these studies, interest yielded a significant effect on deeper comprehension. In a later study, interest affected the recall of total idea units and central ideas as well as the coherence of recall (Schiefele & Krapp, 1991). These findings were consistent with other studies (e.g., Alexander et al., 1994; Entin & Klare, 1985).

Self-concept affects procedural knowledge positively. A stronger academic self-concept was related to more positive motivation (Harter, 1986) and behavior in learning situations (Wigfield & Karpathian, 1991). One example is that learners with a higher self-concept took more coursework (Marsh & Yeung, 1997). This led to better achievement outcomes (Marsh et al., 1999).

However, there are limitations to this finding. The motivational constructs themselves are more innately connected to the knowledge types. Theories on the development of interest focus more on concepts and cognitive components which can be better constituted as declarative knowledge (e.g., Alexander et al., 1994, 1995). Self-concept describes the learner's representation of themselves and their abilities (e.g., Beane & Lipka, 1984; Hattie, 2014;

Shavelson et al., 1976). The term ability is more closely connected to the theoretical representation of procedural knowledge. Therefore, by correlating measurements of interest and self-concept with declarative and procedural knowledge, it is assumable that there are stronger correlations for those constructs that are theoretically more similar. Additionally, according to Tobias (1992, 1994), affective constructs can only indirectly lead to learning by influencing its cognitive processes (Schiefele et al., 1992). This demands more thorough research in the future.

Teachers or instructors should use the interactions found in Study 2 in educational settings by providing supportive offers for learners. Promoting learners' interests and self-concepts aids knowledge acquisition for both types of knowledge. Further research is needed to show how it is possible to apply these supportive offers best in the school curriculum. The instructional methods inspected in Study 3 might yield influences not only on knowledge acquisition and transfer but on motivational factors in students as well (e.g., feedback; Wisniewski et al., 2020). Those motivational factors might again play a role as mediators between prior knowledge and knowledge after learning. Further research is needed to test this idea. Drawing on the results of Study 1 and Study 2, further research could also examine motivational mediators on the cross-sectional and longitudinal relations between conceptual and procedural knowledge in mathematics.

8.1.4 Key Insight 4: The Promotion of Knowledge Transfer Should Start at a Young Age

The school students' age significantly influenced the effectiveness of transfer interventions in Study 3. The younger the children, the more instructional methods in transfer interventions promoted knowledge transfer compared to control interventions. This speaks for the conduction of transfer interventions at kindergarten or preschool. Several theoretical and practical arguments for this add to the finding of Study 3.

From a theoretical perspective, transfer interventions include the acquisition of knowledge that is transferred afterwards. Based on the differentiation in conceptual and procedural knowledge, younger children were found to rely more on procedural knowledge than older children (Hecht et al., 2003; Gelman & Williams, 1998; Halford, 1993; Hiebert & Lefevre, 1986; Schneider et al., 2009). However, we did not find significant differences based on the students' age in Study 1 in mathematics. Conceptual and procedural knowledge were bidirectionally related for both younger and older children. This speaks for the possibility of fostering knowledge as procedural or conceptual knowledge or as both types. Moreover, prior knowledge is among the most influential predictors of later knowledge (e.g., Simonsmeier et al., 2022). Thus, providing a solid knowledge base at a young age prepares children for later knowledge acquisition. These two arguments further underline the relevance of early instructional interventions.

Practically speaking, applying instructional interventions at school is limited in time and resources. There is a curriculum that needs to be considered. However, transfer interventions at preschool or kindergarten are less bound to those restrictions because of a more flexible framework for the general time and content management. This enables teachers to incorporate transfer interventions more freely and target them according to the children's needs.

Several factors should be considered by providing transfer interventions for young children: Young children do not learn effectively through the planned teaching of subjects (Nutbrown, 2006). Instead, the exploration of and engagement with different situations as well as learning through playing are crucial (Nutbrown, 2006; Tizard & Hughes, 2008). Furthermore, it is conceivable that motivational factors mediate prior knowledge and knowledge after learning in young children. This motivation can be fostered, for example, by encouraging children to believe that they can succeed by supporting their autonomy in learning (Deci, Nezlek, et al., 1981; Deci, Schwartz et al., 1981; Koca, 2016; Vallerand et al., 1997).

Some characteristics of young children might mediate the relation found. Gopnik and colleagues (2015) provided two reasons why young children were better at learning unusual abstract causal principles than older ones and adults (Gopnik et al., 2015; Lucas et al., 2014; Seiver et al., 2013): Young children are more flexible in their learning, whereas older learners are more efficient (e.g., Buchsbaum et al., 2012; Chrysikou et al., 2013; Gopnik et al., 2017). This flexibility helps in learning new strategies and topics. Additionally, learning something new, instead of relying on previously inferred general principles, is more difficult for older learners because of their broader prior knowledge (Gopnik et al., 2015). Children have limited cognitive abilities, for example, reduced cognitive control. This aids in learning language, using probabilistic information, and solving causal-reasoning tasks (Thompson-Schill et al., 2009). The faster developing procedural system (Gualtieri & Finn, 2022; Munaka et al., 2011), the increased neural plasticity (e.g., Lenneberg, 1967; Woolley, 2012), and the generally greater exploration (e.g., Gopnik, 2020; Nussenbaum & Hartley, 2019) are additional influential factors.

Transfer interventions are more effective in promoting knowledge transfer for younger than older learners. This result can be used in practical settings by conducting transfer interventions in kindergarten or preschool. Also, the general framework at school, based on limitations in time and resources, speaks for incorporating interventions before entering school. Teachers must consider children's characteristics in learning because there might be mediating influences, for example, the cognitive differences between young and older learners.

8.1.5 Key Insight 5: Current Research Concentrates on Near Knowledge Transfer in School Students

According to Study 3, most primary studies only provided information on near knowledge transfer. It is possible to differentiate knowledge transfer into near and far transfer depending on the similarity between the learning and the transfer situation. This transfer distance can be inspected on several context factors (Barnett & Ceci, 2002): knowledge domain, physical context, temporal context, functional context, social context, and modality. Intervention studies can be ranked on the six context factors, providing a profound basis for inspections. However, the classification of near and far transfer does not imply a clear limit because of the fluent transition of both transfer types (Chi & Wylie, 2014). Therefore, in Study 3, we differentiated near and far transfer based on prior research (Simonsmeier et al., 2022). Assessments of far transfer were mainly found for the factor *modality,* followed by *temporal context*. Far transfer was less often found in the context factors *knowledge domain*, *social context*, *physical context*, and *functional context*.

This limited the results of Study 3 primarily to statements on near knowledge transfer in school students. This result is not surprising: It is assumable that near transfer is more important for learning at school than far transfer because of practical reasons. Students are pressured to learn much content in several school subjects that they need to transfer to related subjects to succeed immediately. A small transfer distance is typical for the knowledge domain (learning and tests on the same topic), physical context (learning and tests at school), functional context (learning and tests are clearly academic), and social context (learning and tests are mostly in class). One could assume that temporal context and modality are the factors that might vary the most in classical school assessments (learning and tests are several days apart; learning and tests differ in their modality). As described above, the number of effect sizes in Study 3 for far transfer in the six context factors confirmed this.

Far transfer is not the focus of interest for the institution *school* for most of the inspected context factors. This makes the results of Study 3 more precise for interventions conducted at school. All examined instructional methods help promote near transfer compared to groups that learned with no or fewer instructional methods. However, the focus on near transfer might not apply to the teachers' goals. They pursue the goal of enabling far transfer in students. Nevertheless, teachers need to adhere to the set framework in terms of restrictions on time and resources.

Future research should investigate the differences between near and far transfer in school students after knowledge interventions. This can be realized by systematically varying the transfer distance on the six context factors according to the Taxonomy of Far Transfer (Barnett & Ceci, 2002). This might yield insights into possible support for teachers and students to enable far knowledge transfer that can be realized at school.

8.1.6 Key Insight 6: Moderator Analyses Can be Enriched with the Use of Latent Class Analyses

Study 3 examined 27 instructional methods as potential moderator variables that might influence the relationship between interventions and knowledge transfer. Most of these methods occurred only sparsely and in varying combinations in the included interventions, making separate moderator analyses negligible. To avoid this problem, we conducted a latent class analysis of the instructional methods. This analysis produced four latent classes based on the responses to the 27 observed indicators (Nylund-Gibson & Choi, 2018). These classes represented unidentified subgroups (Weller et al., 2020) that shared outward characteristics (Hagenaars & McCutcheon, 2002). These class memberships were then used as moderator variables, each with enough effect sizes to constitute reliable results that allow further interpretations. The classes were *Complex Interventions with Multiple Representations and Practice Tasks, Complex Interventions with Demonstrations and Explanations, Self-Explanation,* and *Explanations and Examples of Concepts*.

In contrast to other possible approaches in which instructional methods may be grouped by content-related aspects, the latent class analysis yields several advantages. The process of identifying latent classes of instructional methods is systematic and transparent. The latent class analysis is a model-based approach (Hickendorff et al., 2018; Spurk et al., 2020). With this, it permits a mathematical evaluation of the representation of the included data in the proposed latent class model (Nylund-Gibson & Choi, 2018). The number of classes is systematically identified by class enumeration and fit indices (Muthén, 2003). Because the class level is a latent variable, class membership is based on probability and calculated as a function of its manifest data and the model parameters (Bray et al., 2015; Lanza & Cooper, 2016; Lanza et al., 2013).

There are prior meta-analyses that used latent class analysis in the form of hierarchical Bayesian latent class meta-analyses (e.g., Dendukuri et al., 2012) and latent classes in models of meta-analyses in medicine (e.g., Walter et al., 1999; Yin et al., 2012). However, this approach was seldom done to moderator variables (e.g., Cooper & Lanza, 2014). In this example, subgroups of children were identified by their comprehensive profiles across nine characteristics regarding the beneficial effect of the Head Start Program for these groups (Cooper & Lanza, 2014). The latent moderation approach uncovered variability within subgroups and extended the one-dimensional nature of traditional moderation analyses (e.g., Aiken & West, 1991). With this, more complex structures could be identified (Cooper & Lanza, 2014; Lanza & Cooper, 2016).

This approach makes it possible to simultaneously look at various moderators when there is insufficient evidence for every moderator. Using the four latent classes in Study 3, their differences in promoting knowledge transfer could be investigated. The latent classes varied descriptively, but not significantly, in their results regarding the effectiveness of transfer interventions. This speaks for the general advantage of transfer interventions compared to control interventions. However, a more precise description of included instructional methods would enhance the trustworthiness of latent classes based on these. Additionally, there is a need for further general research to compare single moderators with latent classes as moderating variables in meta-analyses.

8.2 Practical Implications

Several practical implications can be concluded from the results of the three studies. Study 1 showed positive bidirectional predictive relations between conceptual and procedural knowledge in mathematics. This supported the iterative model (Rittle-Johnson et al., 2001): Both knowledge types, conceptual and procedural knowledge, yielded further improvements for the other type and themselves. Acknowledging the significant longitudinal relations found in Study 1, both knowledge types should be incorporated into mathematics education. However, it seems unnecessary to prescribe one of the knowledge types as a starting point because school students show different knowledge profiles (e.g., Hallett et al., 2010; Hecht & Vagi, 2010, 2012). These findings lead to the assumption that it is feasible for teachers to start the instruction with procedural or conceptual knowledge because they both positively influence each other and themselves. Further, combining both knowledge types in the beginning could be even more beneficial for improving knowledge for all knowledge types of school students. Nevertheless, teachers should also consider the characteristics of the sample and the subdomain of mathematics when constructing learning environments.

Teachers need to pay attention to the students' motivation. By increasing their interest and self-concept, teachers can support the success of their students' learning of declarative and procedural knowledge. Educators generally agree on fostering motivation toward learning and school (Anderman & Midgley, 1998; Irvin, 1997). Nevertheless, motivating school students is challenging because of physiological and psychological changes based on developmental processes and learning environment characteristics (Midgley, 1993). Such characteristics include task structure, evaluation strategies, and the quality of teacher-student relationships (Eccles & Midgley, 1989). Teachers work with several heterogeneous groups of students in a short amount of time. This makes it challenging to maintain high motivation for every student. Teaching methods should be inspected regarding their influence on motivation. Such methods to increase motivation in school students reach from whole learning arrangements (e.g., inquiry-based learning; Cairns & Areepattamannil, 2019; Srisawasdi & Panjaburee, 2019) to individual methods in regular school settings, for example, incorporating the identification with academics to support the students' self-concept (Osborne & Jones, 2011). Furthermore, multimedia learning settings are purported to influence motivational factors positively. Examples are augmenting reality game-based learning (Pellas et al., 2019) and immersive virtual reality (Di Natale et al., 2020). By applying such methods, students can profit from the positive influence of motivational factors on their knowledge after learning.

Transfer interventions for school students should incorporate characteristics of the sample and the setting. The age group, the school, and the class level of both groups as well as the mean age of the control groups affected this result. Transfer interventions were generally more effective for younger than older students compared to control interventions. As early as in kindergarten, instructional interventions promoting knowledge provide students with an accurate knowledge basis for further education. Besides other recent studies (e.g., Simonsmeier et al., 2022), Studies 1 and 2 also showed that prior knowledge affected later knowledge. These findings underline the importance of interventions for young students that equip them with accurate knowledge. Adding to this, the school level and the class of the experimental and comparison groups were significant moderators on the advantage of transfer interventions. The higher the school level and class, the smaller the effectiveness of interventions with instructional methods compared to control interventions. This complies with previous findings: Several meta-analyses found small-group reading interventions more effective for younger than older students (Hall & Burns, 2018; Scammacca et al., 2007, 2015). However, Suggate (2016) found similar effects for reading comprehension interventions pre- and at school. The setting of the intervention and the success of implementation checks were also essential moderators. Interventions conducted in realistic settings with successful implementation checks on the learned content positively influenced the advantage of transfer interventions. This contradicts findings on the generally stronger effects of laboratory research (e.g., Vanhove & Harms, 2015). However, it adds to previous results on transfer in the work context, in which trainings in the field and lab had similar effects on post-training knowledge (Blume et al., 2010). The effects found in our meta-analysis were independent of the number of methods used and the number and duration of sessions. Short sessions with fewer instructional methods are as effective as longer sessions with more instructional methods. This underlines the possibility of including transfer interventions in the school curriculum. Teachers and schools should conduct transfer interventions to equip their students with the knowledge they can later transfer to other topics and situations.

8.3 Recommendations for Future Research

The results of this dissertation yield several recommendations for future research. These recommendations can be categorized as methodological recommendations (Chapter 8.3.1) and content-related recommendations (Chapter 8.3.2). Methodological recommendations mainly concentrate on further developing the one-stage MASEM, using a MASEM in mediation analyses, more elaborate models than the cross-lagged panel model, and conducting latent classes as representational moderators. Content-related recommendations comprise the definition of conceptual and procedural knowledge in knowledge domains other than mathematics, the closer inspection of motivational variables as well as the inclusion of more person-related and environmental variables in the relationship between prior knowledge and learning, as well as the integration of definitions of knowledge transfer and instructional methods.

8.3.1 Methodological Recommendations

As mentioned earlier, including structural equation models in meta-analytic investigations in the form of Meta-Analytic Structural Equation Models (MASEM) comprises several advantages. However, there are still possible improvements to the method of the one-stage MASEM. Two examples are described in the following: Up to this point, it is not possible to analyze dependent effect sizes except by choosing one of the dependent effect sizes randomly, choosing one justifiably, or averaging them on the sample level. There are methods for dealing with dependent effect sizes in other meta-analytical models (e.g., robust variance estimation, Hedges et al., 2010). Such methods are also conceivable for one-stage MASEM. Second, examining only two moderator levels simultaneously in categorical moderator analyses is possible. By examining more than two moderator levels in a moderator analysis, the relative importance of each level can be inspected. These improvements are important to gain a better insight into the relations between conceptual and procedural knowledge. Future research should include re-analyzing the model of Study 1 concerning the number of dependent effect sizes and levels of the categorical moderators.

Study 2 showed mediating influences of motivational factors on the relation between prior knowledge and knowledge after learning. Traditionally, this mediator model was examined by analyzing the correlations of all three relations according to the product-of-coefficients strategy (Baron & Kenny, 1986). In future research, conducting a MASEM for this analysis could improve the insights the results provide. Constructing the model as a structural equation model makes it possible to control usual problems, such as missing correlations (Cheung & Cheung, 2016). Additionally, by conducting a MASEM, it is possible to analyze moderating effects on each path coefficient of the whole SEM instead of separate analyses of individual paths of the mediation model. Re-analyzing the model of Study 2 using a MASEM could yield more insights into the importance of motivational mediators and the possible impact of other moderator variables on this mediation.

Another methodological recommendation refers to the use of cross-lagged panel models. As outlined in Study 1, other, more elaborate models yield several advantages (e.g., Random-Intercept Cross-lagged Panel Model; Mund & Nestler, 2019). One advantage is the possibility of distinguishing between-person and within-person variance (Orth et al., 2021). However, the restriction for using such advanced models in primary studies is the availability of at least three measurement points. Researchers in primary research should remember that they must use three or more measurement points to apply elaborate models to the data.

In Study 3, instructional methods were aggregated to latent classes because of the scarcity of effect sizes for individual methods and varying combinations in the included primary studies. These latent classes were then used as moderator variables to produce more profound insights into the influence on the effectiveness of transfer interventions compared to control interventions. This approach could yield advantages for other meta-analyses that include a broad range of moderators with few effect sizes. While there is evidence for the general usage of latent classes in meta-analyses (e.g., Dendukuri et al., 2012; Walter et al., 1999; Yin et al., 2012), moderator variables have rarely been determined by this approach until now (e.g., Cooper & Lanza, 2014). Further research is needed to explain the differences between the influence of single moderators and latent classes as moderating variables.

8.3.2 Content-related Recommendations

In mathematics, researchers usually use the differentiation between conceptual and procedural knowledge and consider these two knowledge types in their studies. It is assumable that both knowledge types also occur in other knowledge domains, such as STEM or literacy. Sénéchal and colleagues (2001) and VanScoy (2019) proposed theoretical models of emergent literacy that incorporate conceptual and procedural knowledge, among other factors. Also, there have been mentions of conceptual and procedural knowledge, for example, in research on technology education (Jones, 1997; McCormick, 1997, 2009), physics (Hestenes, 1987; Taasoobshirazi & Carr, 2008), and biology (Lawson, 2001; Lawson et al., 2000). However, most measurements used in previous studies have no clear distinction between the knowledge types. Aspects of conceptual and procedural knowledge should be identified and classified in tasks and measurements to determine the importance and relations in the respective domain. This could be the basis for further research on longitudinal relations between knowledge types in domains other than mathematics. The results of this research would yield important information for researchers and teachers on the development of knowledge and educational approaches in knowledge domains outside of mathematics.

Motivational factors affect the relation between prior knowledge and learning. This result was shown in Study 2. However, other person-related factors have not been included in this meta-analysis. By integrating 38 meta-analyses, Schneider and Preckel (2017) identified several student variables related to academic achievement. Next to intelligence, prior achievement, and strategies, variables of motivation and personality were important. Besides the motivational factors described in Study 2, the personality factors conscientiousness and test anxiety yielded large effect sizes (Schneider & Preckel, 2017). Such personality variables should be considered as potential mediators when researching knowledge acquisition as well.

The Motivation-Achievement Cycle (Vu et al., 2022) purports the additional impact of social factors. These social factors themselves are not part of the motivation-achievement interactions. However, they affect a person's expectancies and values. Examples are culture or verbal persuasion (Vu et al., 2022). By influencing learners' motivation, these factors could indirectly impact the relation between prior and later knowledge. Further research should include social factors as possible moderators on the mediating effects of motivation.

Besides including other person-related and social factors, the inspected motivational variables could be examined more closely concerning knowledge types. As mentioned earlier, motivational variables, for example, interest and self-concept, are operationalized in measurements without respect to different knowledge types. Therefore, the finding that the construct of interest yielded more influence on declarative knowledge and self-concept on procedural knowledge could be partially attributed to the conceptualizations of the assessments used in the primary studies. Further research could concentrate on the development of measurements concerning the assessed knowledge type.

As presented in Chapter 2.3, the theoretical background of knowledge transfer is diverse and contradictory. There are many theories with different opinions on the reasons for knowledge transfer developed in the last centuries. Researchers in the domain of knowledge transfer should put more effort into unitizing the field of transfer theories in future research. The diversity of transfer theories makes conducting studies in this research area difficult. By constructing an inclusive theory of knowledge transfer, the primary studies could also be better compared to the effectiveness of knowledge transfer. Acknowledging that this is an unrealistic plea, knowledge transfer should be more clearly defined and operationalized in future studies.

This would allow more profound secondary analyses that better capture the kind of transfer examined in the primary study.

Last, Study 3 showed that instructional methods were not well-defined in the research domain of knowledge transfer. By defining these methods more precisely, comparing studies more efficiently with each other would be possible. With this, instructional methods and their impact on knowledge transfer can be better inspected. Several instructional methods from the primary studies were included in the meta-analysis. However, there are plenty more instructional methods that are even less well-defined. It would be essential to define and examine these methods in future studies.

8.4 Conclusion

The three studies in this dissertation encompass three critical processes in the research domains of knowledge acquisition and transfer. The results demonstrated bidirectional predictive relations between conceptual and procedural knowledge in mathematics. Furthermore, motivational factors, for example, interest and self-concept, partially mediated the relation between prior knowledge and knowledge after learning. Lastly, instructional methods in transfer interventions helped promote knowledge transfer in school students compared to control interventions. These results enhance the understanding of how knowledge types are interrelated and give insights into person-related and environmental variables that affect knowledge and transfer.

Using secondary analyses did justice to the broadness of the inspected processes. Each of the three processes has its dedicated research area. By systematically including primary studies in these research areas, the informative values of the meta-analyses were more prominent than in ordinary primary studies. With this advantage, statements based on the results can be more general.

The results of this dissertation demonstrate that knowledge types are bidirectionally interrelated and knowledge acquisition, as well as transfer, are influenced by third variables.

This underlines the dynamic processes that are affected by person-related and environmental factors. A better understanding of this dynamic is feasible by including these factors in the analysis of academic achievement to advance knowledge acquisition.

This dissertation examined crucial knowledge and transfer processes by using meta-analytical models. Each of the three meta-analyses yielded insights into an essential field of educational research. The results of the meta-analyses regarding knowledge acquisition and knowledge transfer are important not only for further research but for practitioners in educational settings as well.

References

Abelson, R. P. (1971). The structure of belief systems. In R. C. Schank & K. M. Colby (Eds.), *Computer models of thought and language* (pp. 287–339), Freeman.

Achmetli, K., Schukajlow, S., & Rakoczy, K. (2019). Multiple Solutions for real-world problems, experience of competence and students' procedural and conceptual knowledge. *International Journal of Science and Mathematics Educations*, *17*(8), 1605–1625. https://doi.org/10.1007/s10763-018-9936-5

Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological Bulletin*, *101*(1), 41–74. https://doi.org/10.1037/0033-2909.101.1.41

Agler R., & De Boeck P. (2017). On the Interpretation and Use of Mediation: Multiple Perspectives on Mediation Analysis. *Frontiers in Psychology*, 8, Article 1984. http://doi.org/10.3389/fpsyg.2017.01984

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. SAGE.

Ainley, M., Hidi, S., & Berndorff, D. (2002). Interest, learning, and the psychological processes that mediate their relationship. *Journal of Educational Psychology*, *94*(3), 545–561. https://doi.org/10.1037/0022-0663.94.3.545

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, *16*(3), 183–198. https://doi.org/10.1016/j.learninstruc.2006.03.001

Ainsworth, S., & Burcham, S. (2007). The impact of text coherence on learning by self-explanation. *Learning and Instruction, 17*(3), 286–303.

Alexander, P. A., Jetton, T. L., & Kulikowich, J. M. (1995). Interrelationship of knowledge, interest, and recall: Assessing a model of domain learning. *Journal of Educational Psychology*, *87*(4), 559–575. https://doi.org/10.1037/0022-0663.87.4.559

Alexander, P. A., Kulikowich, J. M., & Schulze, S. K. (1994). How subject-matter knowledge affects recall and interest. *American Educational Research Journal*, *31*(2), 313–337. https://doi.org/10.2307/1163312

Alexander, P. A., & Murphy, P. (1998). Profiling the differences in students' knowledge, interest, and strategic processing. *Journal of Educational Psychology*, *90*(3), 435–447. https://doi.org/10.1037/0022-0663.90.3.435

Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. *Review of Educational Research*, *61*(3), 315–343. https://doi.org/10.3102/00346543061003315

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist*, *48*(2), 87–113. https://doi.org/10.1080/00461520.2013.775712

Alibali, M. W., & DiRusso, A. A. (1999). The function of gesture in learning to count: More than keeping track. *Cognitive Development*, *14*(1), 37–56. https://doi.org/10.1016/s0885-2014(99)80017-3

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.

An, Y. (2021). A history of instructional media, instructional design, and theories. *International Journal of Technology in Education (IJTE),* *4*(1), 1–21. https://doi.org/10.46328/ijte.35

Anderman, L. H., & Midgley, C. (1998). Motivation and middle school students. *perspective*, *26*(3/4), 325–346.

Anderson, J. R. (1981). Effects of prior knowledge on memory for new information. *Memory & Cognition*, *9*(3), 237–246. https://doi.org/10.3758/bf03196958

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*(4), 369–406. https://doi.org/10.1037/0033-295x.89.4.369

Anderson, J. R. (1983a). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*(3), 261–295. https://doi.org/10.1016/s0022-5371(83)90201-3

Anderson, J. R. (1983b). *The architecture of cognition*. Harvard University.

Anderson, J. R. (1987a). Methodologies for studying human knowledge. *Behavioral and Brain Sciences*, *10*(3), 467–477. https://doi.org/10.1017/s0140525x00023554

Anderson, J. R. (1987b). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review*, *94*, 192–210.

Anderson, J. R. (1993). Production systems and the ACT-R theory. In J. R. Anderson (Ed.), *Rules of the mind* (pp. 17–44). Psychology Press.

Anderson, J. R., Corbett, A. T., Koedinger, K., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences, 4*(2), 167–207. https://doi.org/10.1207/s15327809jls0402_2

Anderson, J. R., & Singley, M. K. (1993). The identical elements theory of transfer. In J. R. Anderson & F. S. Bellezza (Eds.), *Rules of the Mind*, (pp. 183–204). Psychology Press.

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning*, *teaching*, *and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

Arens, A. K., & Hasselhorn, M. (2015). Differentiation of competence and affect self-perceptions in elemantary school students: Extending empirical evidence. *European Journal of Psychology of Education*, *30*(4), 405–419. https://doi.org/10.1007/s10212-015-0247-8

Arens, A. K., Jansen, M., Preckel, F., Schmidt, I., & Brunner, M. (2021). The structure of academic self-concept: A methodological review and empirical illustration of central models. *Review of Educational Research*, *91*(1), 34–72. https://doi.org/10.3102/0034654320972186

Arya, S., Kaji, A. H., & Boermeester, M. A. (2021). PRISMA Reporting guidelines for meta-analyses and systematic reviews. *JAMA Surgery, 156*(8), 789–790. https://doi.org/10.1001/jamasurg.2021.0546

Assink, M., & Wibbelink, C. J. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology*, *12*(3), 154–174. http://dx.doi.org/10.20982/tqmp.12.3.p154

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research*, *70*(2), 181–214.

Ausubel, D. P. (1978). In defense of advance organizers: A reply to the critics. *Review of Educational Research*, *48*(2), 251–257. https://doi.org/10.2307/1170083

Ayer, A. J. (2009). The right to be true. In R. Neta & D. Pritchard (Eds.), *Arguing about knowledge* (pp. 11–13). Routledge.

Aytug, Z. G., Rothstein, H. R., Zhou, W., & Kern, M. C. (2012). Revealed or concealed? Transparency of procedures, decisions, and judgment calls in meta-analyses. *Organizational Research Methods*, *15*(1), 103–133. https://doi.org/10.1177/1094428111403495

Bailey, D. H., Hansen, N., & Jordan, N. C. (2017). The codevelopment of children's fraction arithmetic skill and fraction magnitude understanding. *Journal of Educational Psychology*, *109*(4), 509–519. https://doi.org/10.1037/edu0000152

Bailey, D. H., Hoard, M. K., Nugent, L., & Geary, D. C. (2012). Competence with fractions predicts gains in mathematics achievement. *Journal of Experimental Child Psychology*, *113*(3), 447–455. https://doi.org/10.1016/j.jecp.2012.06.004

Baldwin, R. S., Peleg-Bruckner, Z., & McClintock, A. H. (1985). Effects of topic interest and prior knowledge on reading comprehension. *Reading Research Quarterly*, *20*(4), 497–504. https://doi.org/10.2307/747856

Baldwin, T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology*, *41*(1), 63–105. https://doi.org/10.1111/j.1744-6570.1988.tb00632.x

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. https://doi.org/10.1037/0033-295x.84.2.191

Bandura, A. (1981). Self-referent thought: A developmental analysis of self-efficacy. In J. H. Flavell & l. Ross (Eds.), *Social cognitive development: Frontiers and possible futures* (pp. 200–239). Cambridge University Press.

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

Bandura, A. (1986b). The explanatory and predictive scope of self-efficacy theory. *Journal of Social and Clinical Psychology*, *4*(3), 359–373. https://doi.org/10.1521/jscp.1986.4.3.359

Bandura, A. (1995). *Self-efficacy in changing societies*. Cambridge University Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.

Barlow, H. (1972). Single units and sensation: A neuron doctrine for perceptual psychology? *Perception*, *1*(4), 371–394. https://doi.org/10.1068/p010371

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

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182.

Baroody, A. J. (2003). The development of adaptive expertise and flexibility: The integration of conceptual and procedural knowledge. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills* (pp. 1–34). Routledge.

Baroody, A. J., Feil, Y., & Johnson, A. R. (2007). Research commentary: A rejoinder foregrounding procedural knowledge. *Journal for Research in Mathematics Education*, *38*(2), 115–131. https://doi.org/10.2307/30034953

Baroody, A. J., & Gannon, K. E. (1984). The development of the commutativity principle and economical addition strategies. *Cognition and Instruction*, *1*(3), 321–339. https://doi.org/10.1207/s1532690xci0103_3

Baroody, A. J., & Ginsburg, H. P. (1986). The relationship between initial meaningful and mechanical knowledge of arithmetic. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 75–112). Lawrence Erlbaum.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577–660. https://doi.org/10.1017/s0140525x99002149

Bauer, D. J. (2017). A more general model for testing measurement invariance and differential item functioning. *Psychological Methods*, *22*(3), 507–526. https://doi.org/10.1037/met0000077

Bauer, P. J. (2006). Event memory. In D. Kuhn, R. S. Siegler, W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Cognition*, *perception*, *and language* (pp. 373–425), John Wiley & Sons, Inc.

Baytelman, A., Iordanou, K., & Constantinou, C. P. (2020). Epistemic beliefs and prior knowledge as predictors of the construction of different types of arguments on socioscientific issues. *Journal of Research in Science Teaching*, *57*(8), 1199–1227. https://doi.org/10.1002/tea.21627

Beane, J. A., & Lipka, R. P. (1984). *Self-Concept*, *self-esteem and the curriculum*. Allyn and Bacon, Inc.

Bein, O., Livneh, N., Reggev, N., Gilead, M., Goshen-Gottstein, Y., & Maril, A. (2015).

Delineating the effect of semantic congruency on episodic memory: The role of integration and relatedness. *PloS one*, *10*(2), Article e0115624. https://doi.org/10.1371/journal.pone.0115624

Belenky, D. M., & Nokes-Malach, T. J. (2012). Motivation and transfer: The role of mastery-approach goals in preparation for future learning. *The Journal of the Learning Sciences*, *21*(3), 399–432. https://doi.org/10.1080/10508406.2011.651232

Belland, B. R. (2017). *Instructional scaffolding in STEM education: Strategies and efficacy evidence*. Springer Nature.

Bellezza, F. S. (1981). Mnemonic devices: Classification, characteristics, and criteria. *Review of Educational Research*, *51*(2), 247–275.

Bempeni, M., Poulopoulou, S., Tsiplaki, I., & Vamvakoussi, X. (2018). Individual differences in fractions' conceptual and procedural knowledge: what about older students? In E. Bergqvist, M. Österholm, C. Granberg & L. Sumpter (Eds.), *Proceedings of the 42nd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 147–154). PME.

Bempeni, M., & Vamvakoussi, X. (2014). Seventh and ninth graders' conceptual and procedural knowledge of fractions: what changes with school experience? *Paper presented at the 9th International Conference on Conceptual Change*, *Bologna*, *Italy*.

Bempeni, M., & Vamvakoussi, X. (2015). Individual differences in students' knowing and learning about fractions: Evidence from an in-depth qualitative study. *Frontline Learning Research*, *3*, 17–34.

Bensley, D. A., & Lilienfeld, S. O. (2017). Psychological misconceptions: Recent scientific advances and unresolved issues. *Current Directions in Psychological Science*, *26*(4), 377–382. https://doi.org/10.1177/0963721417699026

Bereiter, C. (1990). Aspects of an educational learning theory. *Review of Educational Research*, *60*(4), 603–624.

Berry, D., & Willoughby, M. T. (2017). On the practical interpretability of cross-lagged panel models: Rethinking a developmental workhorse. *Child Development*, *88*(4), 1186–1206. https://doi.org/10.1111/cdev.12660

Blanchette, I., & Dunbar, K. (2000). How analogies are generated: The roles of structural and superficial similarity. *Memory & Cognition*, *28*(1), 108–124. https://doi.org/10.3758/bf03211580

Bloom, B. S. (1956). *Taxonomy of educational objectives*, *handbook I: The cognitive domain*. David McKay Co Inc.

Bloome, D., Power Carter, S., Morton Christian, B., Otto, S., & Shuart-Faris, N. (2005). *Discourse analysis and the study of classroom language and literacy events: A microethnographic perspective*. Lawrence Erlbaum Associates, Inc.

Blöte, A. W., Van der Burg, E., & Klein, A. S. (2001). Students' flexibility in solving two-digit addition and subtraction problems: Instruction effects. *Journal of Educational Psychology*, *93*(3), 627–638. https://doi.org/10.1037/0022-0663.93.3.627

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

Boaler, J. (2002). The development of disciplinary relationships: knowledge, practice and identity in mathematics classrooms. *For the Learning of Mathematics* *22*(1), 42–47.

Bobko, P., & Stone-Romero, E. F. (1998). Meta-analysis may be another useful research tool, but it is not a panacea. *Research in Personnel and Human Resources*, *16*, 359–397.

Bolisani, E., & Bratianu, C. (2018). *Emergent knowledge strategies: The elusive definition of knowledge* (pp. 1–22). Springer, Cham.

Bong, M., & Clark, R. E. (1999). Comparison between self-concept and self-efficacy in academic motivation research. *Educational Psychologist*, *34*(3), 139–153. https://doi.org/10.1207/s15326985ep3403_1

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, *15*(1), 1–40. https://doi.org/10.1023/A:1021302408382

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

Borenstein, M., Higgins, J. P., Hedges, L. V., & Rothstein, H. R. (2017). Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. *Research Synthesis Methods*, *8*(1), 5–18. https://doi.org/10.1002/jrsm.1230

Braithwaite, D. W., & Sprague, L. (2021). Conceptual Knowledge, Procedural Knowledge, and Metacognition in Routine and Nonroutine Problem Solving. *Cognitive Science*, *45*(10), Article e13048. https://doi.org/10.1111/cogs.13048

Branch, R. M., & Kopcha, T. J. (2014). Instructional Design Models. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology* (pp. 77–87). Springer. http://doi.org/10.1007/978-1-4614-3185-5.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 717–726. https://doi.org/10.1016/S0022-5371(72)80006-9

Bransford, J. D., & Schwartz, D. K. (1999). Chapter 3: Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, *24*(1), 61–100. https://doi.org/10.3102/0091732x024001061

Bray, B. C., Lanza, S. T., & Tan, X. (2015). Eliminating bias in classify-analyze approaches for latent class analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *22*, 1–11. https://doi.org/10.1080/10705511.2014.935265

Briars, D. J., & Siegler, R. S. (1984). A featural analysis of preschoolers' counting knowledge. *Developmental Psychology*, *20*(4), 607–618. https://doi.org/10.1037/0012-1649.20.4.607

Britt, M. A., & Aglinskas, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction*, *20*(4), 485–522. https://doi.org/10.1207/S1532690XCI2004_2

Broadbent, D. E. (1989). Lasting representations and temporary processes. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving*. (pp.211–227). Lawrence Erlbaum Associates, Inc.

Brod, G. (2021). Toward an understanding of when prior knowledge helps or hinders learning. *npj Science of Learning*, *6*(1), Article 24. https://doi.org/10.1038/s41539-021-00103-w

Brod, G., & Shing, Y. L. (2019). A boon and a bane: Comparing the effects of prior knowledge on memory across the lifespan. *Developmental Psychology*, *55*(6), 1326–1337. https://doi.org/10.1037/dev0000712

Brown, A. L., & Kane, M. J. (1988). Preschool children can learn to transfer: Learning to learn and learning from example. *Cognitive Psychology*, *20*(4), 493–523. https://doi.org/10.1016/0010-0285(88)90014-x

Brown, J. M., Collins, A. J., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18*(1), 32–42. https://doi.org/10.3102/0013189x018001032

Brown, S. D., Tramayne, S., Hoxha, D., Telander, K., Fan, X., & Lent, R. W. (2008). Social cognitive predictors of college students' academic performance and persistence: A meta analytic path analysis. *Journal of Vocational Behavior*, *72*(3), 298–308. https://doi.org/10.1016/j.jvb.2007.09.003

Brown, T., & Smith, L. (Eds.). (2003). *Reductionism and the development of knowledge*. Psychology Press.

Bruner, J. S. (1966). *Toward a theory of instruction*. Harvard University Press.

Buchsbaum D., Bridgers S., Weisberg D. S., Gopnik A. (2012). The power of possibility: Causal learning, counterfactual reasoning, and pretend play. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367, 2202–2212. https://doi.org/10.1098/rstb.2012.0122

Bücker, S., Nuraydin, S., Simonsmeier, B. A., Schneider, M., & Luhmann, M. (2018). Subjective well-being and academic achievement: A meta-analysis. *Journal of Research in Personality*, *74*, 83–94. https://doi.org/10.1016/j.jrp.2018.02.007

Buehl, M. M., & Alexander, P. A. (2001). Beliefs about schooled knowledge. *Educational Psychology Review*, *13*, 385–418.

Bulpitt, C. (1988). Meta-analysis. *The Lancet*, *332*, 93–94.

Butterfield, E. C., & Nelson, G. A. (1989). Theory and practice of teaching for transfer. *Educational Technology Research and Development*, *37*(3), 5–38. https://doi.org/10.1007/bf02299054

Byrne, B. M. (1995). Academic self-concept: Its structure, measurement, and relation to academic achievement. In B. A. Bracken (Ed.), *Handbook of self-concept: Developmental*, *social*, *and clinical considerations* (pp. 287–316). John Wiley & Sons.

Byrnes, J. P. (1992). The conceptual basis of procedural learning. *Cognitive Development*, *7*(2), 235–257. https://doi.org/10.1016/0885-2014(92)90013-h

Byrnes, J. P., & Wasik, B. A. (1991). Role of conceptual knowledge in mathematical procedural learning. *Developmental Psychology*, *27*(5), 777–786. https://doi.org/10.1037/0012-1649.27.5.777

Cairns, D., & Areepattamannil, S. (2019). Exploring the relations of inquiry-based teaching to science achievement and dispositions in 54 countries. *Research in Science Education*, *49*, 1–23. https://doi.org/10.1007/s11165-017-9639-x

Canobi, K. H. (2009). Concept-procedure interactions in children’s addition and subtraction. *Journal of Experimental Child Psychology*, *102*(2), 131–149. https://doi.org/10.1016/j.jecp.2008.07.008

Canobi, K. H., Reeve, R. A., & Pattison, P. E. (2003). Patterns of knowledge in children's addition. *Developmental Psychology*, *39*(3), 521–534. https://doi.org/10.1037/0012-1649.39.3.521

Carey, S. (1986). Cognitive science and science education. *American Psychologist*, *41*(10), 1123–1130. https://doi.org/10.1037/0003-066x.41.10.1123

Castel, A. D., McCabe, D. P., Roediger III, H. L., & Heitman, J. L. (2007). The dark side of expertise: Domain-specific memory errors. *Psychological Science*, *18*(1), 3–5. https://doi.org/10.1111/j.1467-9280.2007.01838.x

Catrambone, R. (2002). The effects of surface and structural feature matches on the access of story analogs. *Journal of Experimental Psychology: Learning*, *Memory and Cognition*, *28*(2), 318–334. https://doi.org/10.1037/0278-7393.28.2.318

Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology: Learning*, *Memory and Cognition*, *15*(6), 1147–1156. https://doi.org/10.1037/0278-7393.15.6.1147

Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. *Psychological Bulletin*, *140*(4), 980–1008. https://doi.org/10.1037/a0035661

Chaiklin, S., & Lave, J. (Eds.) (1993). *Understanding practice: Perspectives on activity and context*. Cambridge University Press.

Cheng, E., & Hampson, I. N. (2008). Transfer of training: A review and new insights. *International Journal of Management Reviews*, *10*(4), 327–341. https://doi.org/10.1111/j.1468-2370.2007.00230.x

Cheng, P.-Y., & Chiou, W.-B. (2010). Achievement, attribution, self-efficacy and goal setting by accounting undergraduates. *Psychological Reports*, *106*(1), 54–64. https://doi.org/10.2466/PR0.106.1.54-64

Cheung, M. W. L. (2014). Fixed- and random-effects meta-analytic structural equation modeling: Examples and analyses in R. *Behavior Research Methods*, *46*(1), 29–40. https://doi.org/10.3758/s13428-013-0361-y

Cheung, M. W. L. (2015a). {metaSEM}: An R Package for Meta-Analysis using Structural Equation Modeling. *Frontiers in Psychology 5*, Article 1521. https://doi.org/10.3389/fpsyg.2014.01521

Cheung, M. W. L. (2015b). *Meta-analysis: A structural equation modeling approach*. John Wiley & Sons.

Cheung, M. W. L. (2019). A guide to conducting a meta-analysis with non-independent effect sizes. *Neuropsychology Review*, *29*(4), 387–396. https://doi.org/10.1007/s11065-019-09415-6

Cheung, M. W. L. (2021). Synthesizing Indirect Effects in Mediation Models With Meta-Analytic Methods. *Alcohol and Alcoholism 57*(1), 5–15. https://doi.org/10.1093/alcalc/agab044

Cheung, M. W. L., & Chan, W. (2005). Meta-analytic structural equation modeling: a two-stage approach. *Psychological Methods*, *10*(1), 40–64. https://doi.org/10.1037/1082-989X.10.1.40

Cheung, M. W. L., & Cheung, S. F. (2016). Random-effects models for meta-analytic structural equation modeling: review, issues, and illustrations. *Research Synthesis Methods*, *7*(2), 140–155. https://doi.org/10.1002/jrsm.1166

Cheung, M. W. L., & Vijayakumar, R. (2016). A guide to conducting a meta-analysis. *Neuropsychology Review*, *26*, 121–128. https://doi.org/10.1007/s11065-016-9319-z

Cheung, S. F., & Chan, D. K. (2004). Dependent effect sizes in meta-analysis: Incorporating the degree of interdependence. *Journal of Applied Psychology*, *89*(5), 780–791. https://doi.org/10.1037/0021-9010.89.5.780

Chi, M. T. H. (2008). Three Types of Conceptual Change: Belief Revision, Mental Model Transformation, and Categorical Shift. In S. Vosniadou (Ed.), *Handbook of research on conceptual change* (pp. 61–82). Erlbaum.

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*(1), 73–105. https://doi.org/10.1111/j.1756-8765.2008.01005.x

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*(2), 145–182. https://doi.org/10.1016/0364-0213(89)90002-5

Chi, M. T. H., & Ohlsson, S. (2005). Complex declarative learning. In K. J. Holyoak & R. G. Morrison (Eds.), *The cambridge handbook of thinking and reasoning* (pp. 371–399), Cambridge University Press.

Chi, M. T. H., & VanLehn, K. (2012). Seeing deep structure from the interactions of surface features. *Educational Psychologist*, *47*(3), 177–188. https://doi.org/10.1080/00461520.2012.695709

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Choi, N. (2005). Self-efficacy and self-concept as predictors of college students' academic performance. *Psychology in the Schools*, *42*(2), 197–205. https://doi.org/10.1002/pits.20048

Choi, J. I., & Hannafin, M. (1995). Situated cognition and learning environments: Roles, structures, and implications for design. *Educational Technology Research and Development*, 43, 53–69. https://doi.org/10.1007/BF02300472

Chrysikou E. G., Hamilton R., Coslett H., Datta A., Bikson M., Thompson-Schill S. (2013). Noninvasive transcranial direct stimulation over left prefrontal cortex facilitates cognitive flexibility in tool use. *Cognitive Neuroscience*, 4, 81–89. https://doi.org/10.1080/17588928.2013.768221

Claessens, A., & Engel, M. (2013). How important is where you start? Early mathematics knowledge and later school success. *Teachers College Record*, *115*(6), 1–29. https://doi.org/10.1177/016146811311500603

Clark, D. B., & Linn, M. C. (2013). The knowledge integration perspective: Connections across research and education. *International Handbook of Research on Conceptual Change*, *2*, 61–82.

Clark, R. E., & Voogel, A. (1985). Transfer of training principles for instructional design. *Ectj, 33*(2), 113–123. https://doi.org/10.1007/BF02769112

Clark, S. L., & Muthén, B. (2009). Relating latent class analysis results to variables not included in the analysis. *Unpublished paper available at www.statmodel.com/download/relatinglca.pdf*

Clarke, A. M., Cooper, G. M., & Loudon, E. H. (1969). A set to establish equivalence relations in pre-school children. *Journal of Experimental Child Psychology*, *8*(1), 180–189. https://doi.org/10.1016/0022-0965(69)90039-3

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. L. Erlbaum Associates.

Collins, J. J., & Fauser, B. C. (2005). Balancing the strengths of systematic and narrative reviews. *Human Reproduction Update*, *11*(2), 103–104. https://doi.org/10.1093/humupd/dmh058

Cook, R. D., & Weisberg, S. (1982). *Residuals and influence in regression*. Chapman and Hall.

Cooper, B. R., & Lanza, S. T. (2014). Who benefits most from Head Start? Using latent class moderation to examine differential treatment effects. *Child Development*, *85*(6), 2317–2338. https://doi.org/10.1111/cdev.12278

Cooper, H. (2015). *Research synthesis and meta-analysis: A step-by-step approach*. SAGE Publications.

Cooper, H., Hedges, L. V., & Valentine, J. C. (2018). *The handbook of research synthesis and meta-analysis* (3. Ed.). Russell Sage Foundation.

Covington, M. V. (2000). Intrinsic versus extrinsic motivation in schools: A reconciliation. *Current Directions in Psychological Science*, *9*(1), 22–25.

Covington, M. V., & Müeller, K. J. (2001). Intrinsic versus extrinsic motivation: An approach/avoidance reformulation. *Educational Psychology Review*, *13*(2), 157–176. https://doi.org/10.1023/a:1009009219144

Cowan, R., Donlan, C., Shepherd, D.-L., Cole-Fletcher, R., Saxton, M., & Hurry, J. (2011). Basic calculation proficiency and mathematics achievement in elementary school children. *Journal of Educational Psychology*, *103*, 786–803. https://doi.org/10.1037/a0024556.

Craven, R., & Marsh, H. W. (1997). Threats to gifted and talented students' self concepts in the big pond: Research results and educational implications. *Australasian Journal of Gifted Education*, *6*(2), 7–17.

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

Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, *99*(2), 311–325. https://doi.org/10.1037/0022-0663.99.2.311

Cronbach, L. J., & Suppes, P. (1969). *Research for tomorrow's schools: Disciplined inquiry for education.* Macmillan.

Crook, C. (1991). Computers in the zone of proximal development: Implications for evaluation. *Computers & Education*, *17*(1), 81–91. https://doi.org/10.1016/0360-1315(91)90075-3

Crooks, N. M., & Alibali, M. W. (2014). Defining and measuring conceptual knowledge in mathematics. *Developmental Review*, *34*(4), 344–377. https://doi.org/10.1016/j.dr.2014.10.001

Crowley, K., Shrager, J., & Siegler, R. S. (1997). Strategy discovery as competitive negotiation between metacognitive and associative mechanisms. *Developmental Review, 17*(4), 462–489. https://doi.org/10.1006/drev.1997.0442

Day, J. D., & Cordón, L. A. (1993). Static and dynamic measures of ability: An experimental comparison. *Journal of Educational Psychology*, *85*(1), 75–82. https://doi.org/10.1037/0022-0663.85.1.75

Day, S. H., & Goldstone, R. L. (2011). Analogical transfer from a simulated physical system. *Journal of Experimental Psychology: Learning*, *Memory and Cognition*, *37*(3), 551–567. https://doi.org/10.1037/a0022333

Day, S. H., & Goldstone, R. L. (2012). The import of knowledge export: Connecting findings and theories of transfer of learning. *Educational Psychologist*, *47*(3), 153–176. https://doi.org/10.1080/00461520.2012.696438

de Jong, T., & Ferguson-Hessler, M. G. (1996). Types and qualities of knowledge. *Educational Psychologist*, *31*(2), 105–113. https://doi.org/10.1207/s15326985ep3102_2

Deci, E. L., Nezlek, J., & Sheinman, L. (1981). Characteristics of the rewarder and intrinsic motivation of the rewardee. *Journal of Personality and Social Psychology, 40*, 1–10.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum.

Deci, E. L., & Ryan, R. M. (2000). The "What" and "Why" of goal pursuit: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Deci, E. L., Schwartz, A. J., Sheinman, L., & Ryan, R. M. (1981). An instrument to assess adults' orientations toward control versus autonomy with children: Reflections on intrinsic motivation and perceived competence. *Journal of Educational Psychology, 73*, 642–650.

DeLoache, J. S. (1991). Symbolic functioning in very young children: Understanding of pictures and models. *Child Development*, *62*(4), 736–752. https://doi.org/10.2307/1131174

DeLoache, J. S. (1995). Early understanding and use of symbols: The model model. *Current Directions in Psychological Science*, *4*(4), 109–113. https://doi.org/10.1111/1467-8721.ep10772408

Dendukuri, N., Schiller, I., Joseph, L., & Pai, M. (2012). Bayesian meta-analysis of the accuracy of a test for tuberculous pleuritis in the absence of a gold standard reference. *Biometrics*, *68*(4), 1285–1293. https://doi.org/10.1111/j.1541-0420.2012.01773.x

Descartes, R. (1997). *Key philosophical writings*. Wordsworth Editions.

Dewey, J. (1916). *Democracy and education*. Free Press.

Di Natale, A. F., Repetto, C., Riva, G., & Villani, D. (2020). Immersive virtual reality in K-12 and higher education: A 10-year systematic review of empirical research. *British Journal of Educational Technology*, *51*(6), 2006–2033. https://doi.org/10.1111/bjet.13030

Dickersin, K. (1990). The existence of publication bias and risk Factors for its occurence. *Journal of the American Medical Association*, *263*(10), 1385–1389. https://doi.org/10.1001/jama.1990.03440100097014

Dieckmann, N. F., Slovic, P., & Peters, E. M. (2009). The use of narrative evidence and explicit likelihood by decisionmakers varying in numeracy. *Risk Analysis: An International Journal*, *29*(10), 1473–1488. https://doi.org/10.1111/j.1539-6924.2009.01279.x

diSessa, A. A. (2006). A history of conceptual change research: Threads and fault lines. In Sawyer R. K. (Ed.), *The cambridge handbook of the learning sciences* (pp. 265–281), Cambridge University Press.

Dochy, F. J. R. C. (1994). *Prior knowledge and learning*. Springer.

Dochy, F. J. R. C., & Alexander, P. A. (1995). Mapping prior knowledge: A framework for discussion among researchers. *European Journal of Psychology of Education*, *10*(3), 225–242. https://doi.org/10.1007/BF03172918

Donlan, C., Cowan, R. A., Newton, E. J., & Lloyd, D. (2007). The role of language in mathematical development: Evidence from children with specific language impairments. *Cognition*, *103*(1), 23–33. https://doi.org/10.1016/j.cognition.2006.02.007

Dreistadt, R. (1969). The use of analogies and incubation in obtaining insights in creative problem solving. *The Journal of Psychology*, *71*(2), 159–175. https://doi.org/10.1080/00223980.1969.10543082

Dreyfus, T. (1991). Advanced mathematical thinking processes. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 25–41). Springer Dordrecht.

Durnin, J. H., Perrone, A. E., & MacKay, L. (1997). Teaching problem solving processes in elementary school mathematics. *Journal of Structural Learning & Intelligent Systems*, *13*(1), 53–69.

Easterbrook, P. J., Gopalan, R., Berlin, J. A., & Matthews, D. R. (1991). Publication bias in clinical research. *The Lancet*, *337*(8746), 867–872. https://doi.org/10.1016/0140-6736(91)90201-Y

Eaton, J., Anderson, C. E., & Smith, E. C. (1984). Students' misconceptions interfere with science learning: Case studies of fifth-grade students. *Elementary School Journal*, *84*(4), 365–379. https://doi.org/10.1086/461370

Eccles, J. S., & Midgley, C. (1989). Stage-environment fit: Developmentally appropriate classrooms for young adolescents. *Research on Motivation in Education*, *3*(1), 139–186.

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

Egger, M., Smith, G. D., & Minder, C. (1997). Bias in metaanalysis detected by a simple, graphical test. *British Medical Journal*, *315*, 629–634. https://doi.org/10.1136/bmj.315.7109.629

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, *1*(1), 49–98.

Elliot, A. J., McGregor, H. A., & Gable, S. (1999). Achievement goals, study strategies, and exam performace: A mediational analysis. *Journal of Educational Psychology*, *91*(3), 549–563. https://doi.org/10.1037/0022-0663.91.3.549

Ellis, A. B. (2007a). A taxonomy for categorizing generalizations: Generalizing actions and reflection generalizations. *The Journal of the Learning Sciences*, *16*(2), 221–262. https://doi.org/10.1080/10508400701193705

Ellis, A. B. (2007b). Connections between generalizing and justifying: Students' reasoning with linear relationships. *Journal for Research in Mathematics Education, 38*(3), 194–229. https://doi.org/10.2307/30034866

Ellis, A. B. (2011). Generalizing-promoting actions: How classroom collaborations can support students' mathematical generalizations. *Journal for Research in Mathematics Education, 42*(4), 308–345. https://doi.org/10.5951/jresematheduc.42.4.0308

Ellis, H. C. (1965). *The transfer of learning*. Macmillan.

Engle, R. A. (2006). Framing interactions to foster generative learning: A situative explanation of transfer in a community of learners classroom. *The Journal of the Learning Sciences*, *15*(4), 451–498. https://doi.org/10.1207/s15327809jls1504_2

Engle, R. A., Lam, D. P., Meyer, X. S., & Nix, S. A. (2012). How does expansive framing promote transfer? Several proposed explanations and a research agenda for investigating them. *Educational Psychologist*, *47*(3), 215–231. https://doi.org/10.1080/00461520.2012.695678

Engle, R. A., Nguyen, P. H., & Mendelson, A. (2011). The influence of framing on transfer: initial evidence from a tutoring experiment. *Instructional Science*, *39*(5), 603–628. https://doi.org/10.1007/s11251-010-9145-2

Entin, E. B., & Klare G. R. (1985). Relationships of measures of interest, prior knowledge, and readability to comprehension of expository passages. *Advances in Reading/Language Research 3*, 9–38.

Epskamp, S. (2019). *semPlot: Path diagrams and visual analysis of various SEM packages' output*. *R package version 1.1.2*. https://CRAN.R-project.org/package=semPlot

Ericsson, K. A. (2003). Exceptional memorizers: Made, not born. *Trends in Cognitive Sciences*, *7*(6), 233–235. https://doi.org/10.1016/s1364-6613(03)00103-7

Fagard, R., Staessen, J. A., & Thijs, L. (1996). Advantages and disadvantages of the meta-analysis approach. *Journal of Hypertension*, *14*(Supplement 2), 9–13. https://doi.org/10.1097/00004872-199609002-00004

Farkas, R. D. (2003). Effects of traditional versus learning-styles instructional methods on middle school students. *The Journal of Educational Research*, *97*(1), 42–51. https://doi.org/10.1080/00220670309596627

Faulkenberry, T. J. (2013). The conceptual/procedural distinction belongs to strategies, not tasks: A comment on Gabriel et al.(2013). *Frontiers in Psychology*, *4*, Article 820. https://doi.org/10.3389/fpsyg.2013.00715

Faulkenberry, T. J., & Pierce, B. H. (2011). Mental representations in fraction comparison. *Experimental Psychology, 58*(6), 480–489. https://doi.org/10.1027/1618-3169/a000116

Fazio, L. K., Barber, S., Rajaram, S., Ornstein, P. A., & Marsh, E. J. (2013). Creating illusions of knowledge: Learning errors that contradict prior knowledge. *Journal of Experimental Psychology: General*, *142*(1), 1–5. https://doi.org/10.1037/a0028649

Feldman, J. M. (1986). *Neural representation of conceptual knowledge*. Rochester University NY Dept of Computer Science.

Fennema, E. H. (1972). Models and mathematics. *The Arithmetic Teacher*, *19*(8), 635–640. https://doi.org/10.5951/AT.19.8.0635

Ferla, J., Valcke, M., & Cai, Y. H. (2009). Academic self-efficacy and academic self-concept: Reconsidering structural relationships. *Learning and Individual Differences*, *19*(4), 499–505. https://doi.org/10.1016/j.lindif.2009.05.004

Fidell, L. S., & Tabachnick, B. G. (2003). Preparatory data analysis. In J. A. Schinka & W. F. Velicer (Eds.), *Handbook of psychology: Research methods in psychology*, Vol. 2, (pp. 115–141). John Wiley & Sons Inc.

Field, A. P., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology*, *63*(3), 665–694. https://doi.org/10.1348/000711010X502733

Fong, C. J., Patall, E. A., Vasquez, A. C., & Stautberg, S. (2019). A meta-analysis of negative feedback on intrinsic motivation. *Educational Psychology Review*, *31*, 121–162. https://doi.org/10.1007/s10648-018-9446-6

Fong, G. T., Evans, M. I., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology*, *18*(3), 253–292. https://doi.org/10.1016/0010-0285(86)90001-0

Forbus, K., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science*, *19*(2), 141–205. https://doi.org/10.1016/0364-0213(95)90016-0

Ford, J. K., & Weissbein, D. A. (1997). Transfer of training: An updated review and analysis. *Performance Improvement Quarterly*, *10*(2), 22–41. https://doi.org/10.1111/j.1937-8327.1997.tb00047.x

Frade, C., Winbourne, P., & Braga, S. M. (2009). A mathematics-science community of practice: Reconceptualising transfer in terms of crossing boundaries. *For the learning of mathematics*, *29*(2), 14–22. https://flm-journal.org/Articles/21CEAFC1625DF527CE1198DD2E56.pdf

Francis, G. (2013). Replication, statistical consistency, and publication bias. *Journal of Mathematical Psychology*, *57*(5), 153–169. https://doi.org/10.1016/j.jmp.2013.02.003

Fries, L., Son, J. Y., Givvin, K. B., & Stigler, J. W. (2021). Practicing connections: A framework to guide instructional design for developing understanding in complex domains. *Educational Psychology Review*, *33*(2), 739–762. https://doi.org/10.1007/s10648-020-09561-x

Fries, S., & Souvignier, E. (2009). Training. In Wild, E., Möller, J. (Eds). *Pädagogische Psychologie* (pp. 405–424). Springer. https://doi.org/10.1007/978-3-642-41291-2_17

Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science*, *18*(3), 233–239. https://doi.org/10.1111/j.1467-9280.2007.01882.x

Fritz M. S., Taylor A. B., & MacKinnon D. P. (2012). Explanation of Two Anomalous Results in Statistical Mediation Analysis. *Multivariate Behavioral Research*, *47*(1), 61–87. https://doi.org/10.1080/00273171.2012.640596

Frye, D., Braisby, N., Lowe, J., Maroudas, C., & Nicholls, J. (1989). Young children's understanding of counting and cardinality. *Child Development*, *60*(5), 1158–1171.

Fyfe, E. R., McNeil, N. M., Son, J., & Goldstone, R. L. (2014). Concreteness fading in mathematics and science instruction: A systematic review. *Educational Psychology Review*, *26*(1), 9–25. https://doi.org/10.1007/s10648-014-9249-3

Gabriel, F., Coche, F., Szucs, D., Carette, V., Rey, B., & Content, A. (2013). A componential view of children's difficulties in learning fractions. *Frontiers in Psychology, 4*, Article 715. https://doi.org/10.3389/fpsyg.2013.00715

Gagné, R. M. (1965). *The conditions of learning* (1st Ed.). Holt, Rinehart & Winston.

Gärdenfors, P. (2017). Demonstration and pantomime in the evolution of teaching. *Frontiers in Psychology*, *8*, Article 415. https://doi.org/10.3389/fpsyg.2017.00415

Gary, M. S., Wood, R. J., & Pillinger, T. (2012). Enhancing mental models, analogical transfer, and performance in strategic decision making. *Strategic Management Journal*, *33*(11), 1229–1246. https://doi.org/10.1002/smj.1979

Gauthier, I., Tarr, M., & Bub, D. (2010). *Perceptual expertise: Bridging brain and behavior*. Oxford University Press.

Geary, D. C. (1994). Developing arithmetical skills. In D. C. Geary (Ed.), *Children's mathematical development: Research and practical applications* (pp. 37–93). American Psychological Association. https://doi.org/10.1037/10163-002

Geary, D. C. (1995). Reflections of evolution and culture in children's cognition: Implications for mathematical development and instruction. *American Psychologist, 50*(1), 24–37.

Gegenfurtner, A., Lehtinen, E., & Säljo, R. (2011). Expertise Differences in the Comprehension of visualization: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review*, *23*(4), 523–552. https://doi.org/10.1007/s10648-011-9174-7

Gelman, R. (1993). A rational-constructivist account of early learning about numbers and objects. In D. L. Medin (Ed.), *The psychology of learning and motivation* (pp. 61–96), Academic Press.

Gelman, R., & Williams, E. M. (1998). Enabling constraints for cognitive development and learning: Domain specificity and epigenesis. In D. Kuhn & R. S. Siegler (Eds.), *Handbook of child psychology: Vol. 2. cognition*, *perception*, *and language* (5th Ed., pp. 575–630), Wiley.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*(2), 155–170. https://doi.org/10.1207/s15516709cog0702_3

Gentner, D. (2005). The development of relational category knowledge. In L. Gershkoff-Stowe & D. H. Rakison (Eds.), *Building object categories in developmental time* (pp. 263–294). Psychology Press.

Gentner, D., & Colhoun, J. (2010). Analogical processes in human thinking and learning. In B. Glatzeder, V. Goel & A. Müller (Eds.), *Towards a theory of thinking* (pp. 35–48). Springer eBooks https://doi.org/10.1007/978-3-642-03129-8_3

Gentner, D., & Gentner, D. R. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Centner & A. L. Stevens (Eds.), *Mental models* (pp. 99–129). Erlbaum.

Gentner, D., Holyoak, K. J., & Kokinov, B. N. (2001). *The analogical mind: Perspectives from cognitive science*. MIT Press.

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, *95*(2), 393–408. https://doi.org/10.1037/0022-0663.95.2.393

Gettier, E. (2009). Is justified true belief knowledge? In R. Neta & D. Pritchard (Eds.), *Arguing about knowledge*. (pp. 14–15). Routledge.

Gibson, J. J. (1977). The theory of affordances. In R. E. Shaw & J. Bransford (Eds.), *Perceiving*, *acting*, *and knowing* (pp. 67–82). Erlbaum.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology*, *12*(3), 306–355. https://doi.org/10.1016/0010-0285(80)90013-4

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*(1), 1–38. https://doi.org/10.1016/0010-0285(83)90002-6

Gick, M. L., & Holyoak, K. J. (1987). The cognitive basis of knowledge transfer. In S.M. Cormier, S.D. Hagman (Eds), *Transfer of learning: Contemporary research and applications* (pp. 9–46). Elsevier eBooks. https://doi.org/10.1016/b978-0-12-188950-0.50008-4

Gist, M. E., & Mitchell, T. R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management Review*, *17*(2), 183–211. https://doi.org/10.5465/amr.1992.4279530

Glaser, R. (1962). Psychology and instructional technology. In R. Glaser (Ed.), *Training research and education* (pp. 1–30). University of Pittsburgh Press.

Gobet, F. (2005). Chunking models of expertise: implications for education. *Applied Cognitive Psychology*, *19*(2), 183–204. https://doi.org/10.1002/acp.1110

Goldin-Meadow, S. (2005). *Hearing gesture: How our hands help us think*. Harvard University Press.

Goldin-Meadow, S., Alibali, M. W., & Church, R. B. (1993). Transitions in concept acquisition: using the hand to read the mind. *Psychological Review*, *100*, 279–297. https://doi.org/10.1037//0033-295X.100.2.279.

Goldstone, R. L., & Sakamoto, Y. (2003). The transfer of abstract principles governing complex adaptive systems. *Cognitive Psychology*, *46*(4), 414–466. https://doi.org/10.1016/s0010-0285(02)00519-4

Gollob, H. F., & Reichardt, C. S. (1987). Taking account of time lags in causal models. *Child Development*, *58*(1), 80–92.

Gopnik, A. (2020). Childhood as a solution to explore–exploit tensions. *Philosophical*

Transactions of the Royal Society B: Biological Sciences, 375(1803), Article 20190502. https://doi.org/10.1098/rstb.2019.0502

Gopnik, A., Griffiths, T. L., & Lucas, C. G. (2015). When younger learners can be better (or at least more open-minded) than older ones. *Current Directions in Psychological Science*, *24*(2), 87–92. https://doi.org/10.1177/0963721414556653

Gorham, G. (2002). Descartes on the innateness of all ideas. *Canadian Journal of Philosophy, 32*(3), 355–388. https://doi.org/10.1080/00455091.2002.10716523

Graesser, A. C., & Person, N. K. (1994). Question asking during tutoring. *American Educational Research Journal*, *31*(1), 104–137. https://doi.org/10.3102/00028312031001104

Greco, T., Zangrillo, A., Biondi-Zoccai, G., & Landoni, G. (2013). Meta-analysis: pitfalls and hints. *Heart*, *Lung and Vessels*, *5*(4), 219–225.

Greeno, J. G. (1997). On claims that answer the wrong questions. *Educational Researcher*, *26*(1), 5–17. https://doi.org/10.3102/0013189x026001005

Greeno, J. G., Moore, J. L., & Smith, D. R. (1993). Transfer of situated learning. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence*, *cognition*, *and instruction* (pp. 99–167). Ablex Publishing.

Greve, A., Cooper, E., Tibon, R., & Henson, R. N. (2019). Knowledge is power: Prior Knowledge aids memory for both congruent and incongruent events, but in different ways. *Journal of Experimental Psychology*, *148*(2), 325–341. https://dx.doi.org/10.1037/xge0000498

Grossman, R., & Salas, E. (2011). The transfer of training: what really matters. *International Journal of Training and Development*, *15*(2), 103–120. https://doi.org/10.1111/j.1468-2419.2011.00373.x

Gualtieri, S., & Finn, A. S. (2022). The sweet spot: When children's developing abilities, brains, and knowledge make them better learners than adults. *Perspectives on*

Psychological Science, *17*(5), 1322–1338. https://doi.org/10.1177/17456916211045971

Guay, F., Marsh, H. W., & Boivin, M. (2003). Academic self-concept and academic achievement: Developmental perspectives on their causal ordering. *Journal of Educational Psychology*, *95*(1), 124–136. https://doi.org/10.1037/0022-0663.95.1.124

Guay, F., Ratelle, C. F., Roy, A., & Litalien, D. (2010). Academic self-concept, autonomous academic motivation, and academic achievement: Mediating and additive effects. *Learning and Individual Differences*, *20*(6), 644–653. https://doi.org/10.1016/j.lindif.2010.08.001

Guberman, S. R., & Greenfield, P. M. (1991). Learning and transfer in everyday cognition. *Cognitive Development*, *6*(3), 233–260. https://doi.org/10.1016/0885-2014(91)90038-F

Gulley, N. (2013). *Plato's theory of knowledge*. Routledge.

Gunderson, E. A., & Hildebrand, L. (2021). Relations among spatial skills, number line estimation, and exact and approximate calculation in young children. *Journal of Experimental Child Psychology*, *212*, Article 105251. https://doi.org/10.1016/j.jecp.2021.105251

Gustafson, K. L., & Branch, R. M. (2002). What is Instructional Design? In R. A. Reiser & J. A. Dempsey (Eds.), *Trends and issues in instructional design and technology*. Merrill/Prentice Hall. http://doi.org/10.1016/B978-0-12-386531-1.00002-8

Gysin, D., & Brovelli, D. (2021). Use of knowledge pieces and context features during the transfer process in physics tasks. *International Journal of Science Education*, *43*(13), 2108–2126. https://doi.org/10.1080/09500693.2021.1952334

Haapasalo, L., & Kadijevich, D. (2000). Two types of mathematical knowledge and their relation. *Journal für Mathematikdidaktik*, *21*(2), 139–157. https://doi.org/10.1007/bf03338914

Hagenaars, J. A., & McCutcheon, A. L. (2002). *Applied latent class analysis*. Cambridge University Press.

Hahn, R. A., & Truman, B. I. (2015). Education improves public health and promotes health equity. *International Journal of Health Services*, *45*(4), 657–678. https://doi.org/10.1177/0020731415585986

Hailikari, T., Katajavuori, N., & Lindblom-Ylänne, S. (2008). The relevance of prior knowledge in learning and instructional design. *The American Journal of Pharmaceutical Education*, *72*(5), Article 113. https://doi.org/10.5688/aj7205113

Halford, G. S. (1993). *Children's understanding: The development of mental models*. Erlbaum.

Hall, M. S., & Burns, M. K. (2018). Meta-analysis of targeted small-group reading interventions. *Journal of School Psychology*, *66*, 54–66. https://doi.org/10.1016/j.jsp.2017.11.002

Hall, W. S. (1989). Reading comprehension. *American Psychologist*, *44*(2), 157–161. https://doi.org/10.1037/0003-066X.44.2.157

Hallett, D., Nunes, T., & Bryant, P. (2010). Individual differences in conceptual and procedural knowledge when learning fractions. *Journal of Educational Psychology*, *102*(2), 395–406. https://doi.org/10.1037/a0017486

Hallett, D., Nunes, T., Bryant, P., & Thorpe, C. M. (2012). Individual differences in conceptual and procedural fraction understanding: The role of abilities and school experience. *Journal of Experimental Child Psychology, 113*(4), 469–486. https://doi.org/10.1016/j.jecp.2012.07.009

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*(1), 102–116. https://doi.org/10.1037/a0038889

Hammer, D., Elby, A., Scherr, R. E., & Redish, E. F. (2005). Resources, framing, and transfer. In J. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 89–119), Information Age.

Hansen, N., Jordan, N. C., Fernandez, E., Siegler, R. S., Fuchs, L., Gersten, R., & Micklos, D. (2015). General and math-specific predictors of sixth-graders' knowledge of fractions. *Cognitive Development*, *35*, 34–49.

Hansen, N., Jordan, N. C., & Rodrigues, J. (2017). Identifying learning difficulties with fractions: A longitudinal study of student growth from third through sixth grade. *Contemporary Educational Psychology*, *50*, 45–59. https://doi.org/10.1016/j.cogdev.2015.02.001

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., Carter, S. M., & Elliot, E. J. (2000). Short-term and long-term consequences of achievement: Predicting continued interest and performance over time. *Journal of Educational Psychology*, *92*, 316–330.

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., & Elliot, A. J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology, 94*(3), 562–575. https://doi.org/10.1037/0022-0663.94.3.562

Harter, S. (1986). Processes underlying the construction, maintenance, and enhancement of self-concept in children. In S. Suls & A. Greenwald (Eds.), *Psychological perspectives of the self* (Vo1. 3, pp. 136–182). Lawrence Erlbaum Associates, Inc.

Hartley, R. T., & Barnden, J. A. (1997). Semantic networks: visualizations of knowledge. *Trends in Cognitive Sciences*, *1*(5), 169–175.

Haskell, R. E. (2001). Chapter 2 - Transfer of Learning: What It Is and Why It's Important. In R. E. Haskell (Ed.), *Transfer of learning. Cognition, instruction, and reasoning* (pp. 23–39). Academic Press. https://doi.org/10.1016/b978-012330595-4/50003-2

Hattie, J. (2014). *Self-concept*. Psychology Press.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, *76*(4), 408–420. https://doi.org/10.1080/03637750903310360

Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach (3rd Ed.)*. The Guilford Press

Hayes, J. R., & Simon, H. A. (1977). Psychological differences among problem isomorphs. In N. J. Castellan, D. B. Pisoni & G. R. Potts (Eds.), *Cognitive theory* (Vol. 2, pp. 21), Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Hecht, S. A., Close, L., & Santisi, M. (2003). Sources of individual differences in fraction skills. *Journal of Experimental Child Psychology*, *86*(4), 277–302. https://doi.org/10.1016/j.jecp.2003.08.003

Hecht, S. A., & Vagi, K. J. (2010). Sources of group and individual differences in emerging fraction skills. *Journal of Educational Psychology*, *102*(4), 843–859. https://doi.org/10.1037/a0019824

Hecht, S. A., & Vagi, K. J. (2012). Patterns of strengths and weaknesses in children's knowledge about fractions. *Journal of Experimental Child Psychology*, *111*(2), 212–229. https://doi.org/10.1016/j.jecp.2011.08.012

Heckhausen, J., & Heckhausen, H. (2018). Motivation and action: Introduction and overview. In J. Heckhausen & H. Heckhausen (Eds.), *Motivation and action* (pp. 1–14). Springer International Publishing. https://doi.org/10.1007/978-3-319-65094-4_1

Hedges, L. V. (1992). Meta-analysis. *Journal of Educational Statistics*, *17*(4), 279–296.

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, *1*(1), 39–65. https://doi.org/10.1002/jrsm.5

Helmke, A., & Van Aken, M. A. (1995). The causal ordering of academic achievement and self-concept of ability during elementary school: A longitudinal study. *Journal of Educational Psychology*, *87*(4), 624–637. https://doi.org/10.1037/0022-0663.87.4.624

Henderson, J. B. (2019). Beyond “active learning”: How the ICAP framework permits more acute examination of the popular peer instruction pedagogy. *Harvard Educational Review*, *89*(4), 611–634. https://doi.org/10.17763/1943-5045-89.4.611

Hennessy, S. (1993). Situated cognition and cognitive apprenticeship: Implications for classroom learning. *Studies in Science Education*, *22*(1), 1–41. https://doi.org/10.1080/03057269308560019

Hershkowitz, R., Ben-Chaim, D., Hoyles, C., Lappan, G., Mitchelmore, M. and Vinner, S. (1989). Psychological aspects of learning geometry. In P. Nesher and J. Kilpatrick (Eds.), *Mathematics and Cognition* (pp. 70–95). University Press, Cambridge.

Hertzog, C., & Nesselroade, J. R. (2003). Assessing Psychological Change in Adulthood: An Overview of Methodological Issues. *Psychology and Aging, 18*(4), 639–657. https://doi.org/10.1037/0882-7974.18.4.639

Hestenes, D. (1987). Toward a modeling theory of physics instruction. *American Journal of Physics*, *55*(5), 440–454.

Hickendorff, M., Edelsbrunner, P. A., McMullen, J., Schneider, M., & Trezise, K. (2018). Informative tools for characterizing individual differences in learning: Latent class, latent profile, and latent transition analysis. *Learning and Individual Differences*, *66*, 4–15. https://doi.org/10.1016/j.lindif.2017.11.001

Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research*, *60*(4), 549–571. https://doi.org/10.3102/00346543060004549

Hidi, S. (1995). A reexamination of the role of attention in learning from text. *Educational Psychological Review*, 7, 323–350.

Hidi, S., & Baird, W. M. (1986). Interestingness? A neglected variable in discourse processing. *Cognitive Science, 10*(2), 179–194. https://doi.org/10.1016/s0364-0213(86)80003-9

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

https://doi.org/10.1207/s15326985ep4102_4

Hidi, S., Renninger, K. A., & Krapp, A. (2004). Interest, a motivational variable that combines affective and cognitive functioning. In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 89–115). Lawrence Erlbaum Associates, Inc.

Hiebert, J. (1992). Mathematical, cognitive, and instructional analyses of decimal fractions. In G. Leinhardt, R. Putnam, & R. Hattrup (Eds.), *Analysis of arithmetic for mathematics teaching* (pp. 283–322). Erlbaum.

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics (*pp. 1–27), Routledge.

Hiebert, J., & Wearne, D. (1996). Instruction, understanding, and skill in multidigit addition and subtraction. *Cognition and Instruction*, *14*(3), 251–283. https://doi.org/10.1207/s1532690xci1403_1

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, *327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Hintzman, D. L. (1976). Repetition and memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 10, pp. 47–91). Academic Press.

Hintzman, D. L. (1986). „Schema abstraction“ in a multiple-trace memory model. *Psychological Review*, *93*(4), 411–428. https://doi.org/10.1037/0033-295x.93.4.411

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, *95*(4), 528–551. https://doi.org/10.1037/0033-295x.95.4.528

Hintzman, D. L., & Block, R. A. (1971). Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology, 88*(3), 297–306. https://doi.org/10.1037/h0030907

Hodges, C. B. (2008). Self-efficacy in the context of online learning environments: A review of the literature and directions for research. *Performance Improvement Quarterly*, *20*(3-4), 7–25. https://doi.org/10.1002/piq.20001

Hogan, K., Nastasi, B. K., & Pressley, M. (1999). Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognition and Instruction*, *17*, 379–432.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). *Induction: Processes of inference*, *learning*, *and discovery*. MIT Press.

Holyoak, K. J. (1985). The pragmatics of analogical transfer. *The Psychology of Learning and Motivation*, *19*, 59–87. https://doi.org/10.1016/S0079-7421(08)60524-1

Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition*, *15*(4), 332–340. https://doi.org/10.3758/bf03197035

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, *13*(3), 295–355. https://doi.org/10.1207/s15516709cog1303_1

Homa, D., Cross, J., Cornell, D., Goldman, D., & Schwartz, S. (1973). Prototype abstraction and classification of new instances as a function of number of instances defining the prototype. *Journal of Experimental Psychology, 101*(1), 116–122. https://doi.org/10.1037/h0035772

Homa, D., Sterling, S., & Trepel, L. (1981). Limitations of exemplar-based generalization and the abstraction of categorical information. *Journal of Experimental Psychology: Human Learning and Memory, 7*(6), 418–439. https://doi.org/10.1037/0278-7393.7.6.418

Honebein, P. C. (2022). The influence of motivational values on instructional designers' values about methods. *Educational technology research and development*, *70*(2), 447–467. https://doi.org/10.1007/s11423-022-10087-y

Honebein, P. C., & Reigeluth, C. M. (2020). The instructional theory framework appears lost. Isn't it time we find it again? *Revista De Educación a Distancia, 20*(64), 1–24. https://doi.org/10.6018/red.405871

Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review*, *17*, 63–84. https://doi.org/10.1016/j.edurev.2015.11.002

Hossack, K. (2007). *The metaphysics of knowledge*. OUP Oxford.

Howard, J. L., Bureau, J., Guay, F., Chong, J. X. Y., & Ryan, R. M. (2021). Student motivation and associated outcomes: A meta-analysis from self-determination theory. *Perspectives on Psychological Science, 16*(6), 1300–1323. https://doi.org/10.1177/1745691620966789

Huber, M. (2020). Mediation Analysis. In K. F. Zimmermann (Ed), *Handbook of labor*, *human resources and population economics* (pp. 1–38). Springer. https://doi.org/10.1007/978-3-319-57365-6_162-1

Hull, C. L. (1943). *Principles of behavior*. Appleton-Century-Crofts.

Hunter, J. E. & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Sage.

Hurrell, D. (2021). Conceptual knowledge or procedural knowledge or conceptual knowledge and procedural knowledge: Why the conjunction is important to teachers. *Australian Journal of Teacher Education (Online)*, *46*(2), 57–71. http://dx.doi.org/10.14221/ajte.2021v46n2.4

Iacobucci, D. (2008). *Mediation analysis (No. 156)*. Sage.

Igartua, J. J., & Hayes, A. F. (2021). Mediation, moderation, and conditional process analysis: Concepts, computations, and some common confusions. *The Spanish Journal of Psychology*, *24*, Article e49.

Inglis, M., & Attridge, N. (2016). *Does mathematical study develop logical thinking? Testing the theory of formal discipline*. World Scientific.

Irvin, J. L. (1997). *What current research says to the middle level practitioner*. National Middle School Association.

Jak, S., & Cheung, M. W. (2020). Meta-analytic structural equation modeling with moderating effects on SEM parameters. *Psychological Methods*, *25*(4), 430–455. https://doi.org/10.1037/met0000245

Jak, S., Li, H., Kolbe, L., de Jonge, H., & Cheung, M. W. (2021). Meta-analytic structural equation modeling made easy: A tutorial and web application for one-stage MASEM. *Research Synthesis Methods*, *12*(5), 590–606. https://doi.org/10.1002/jrsm.1498

Jelsma, O., van Merriënboer, J. J., & Bijlstra, J. P. (1990). The ADAPT design model: Towards instructional control of transfer. *Instructional Science*, *19*(2), 89–120. https://doi.org/10.1007/BF00120698

Johnson, A. M., Reisslein, J., & Reisslein, M. (2014). Representation sequencing in computer-based engineering education. *Computers & Education*, *72*, 249–261. https://doi.org/10.1016/j.compedu.2013.11.010

Johnson, C., & Mayer, R. E. (2010). Applying the self-explanation principle to multimedia learning in a computer-based game-like environment. *Computers in Human Behavior*, *26*(6), 1246–1252. https://doi.org/10.1016/j.chb.2010.03.025

Johnson-Laird, P. N. (2006). *How we reason*. Oxford University Press.

Jones, A. (1997). Recent research in student learning of technological concepts and processes. In M. J. Vries & A. Tamir (Eds.), *Shaping Concepts of Technology* (pp. 83–96). Springer Dordrecht.

Jordan, C. L., Hansen, N., Fuchs, L., Siegler, R., Micklos, D., & Gersten, R. (2013). Developmental predictors of conceptual and procedural knowledge of fractions.

Journal of Experimental Child Psychology, *116*(1), 45–58. https://doi.org/10.1016/j.jecp.2013.02.001

Judd, C. H. (1908). The relation of special training and general intelligence. *Educational Review*, *36*, 28–42. https://ci.nii.ac.jp/naid/10031089770

Judge, T. A., Cable, D. M., Boudreau, J. W., & Bretz Jr, R. D. (1995). An empirical investigation of the predictors of executive career success. *Personnel Psychology*, *48*(3), 485–519. https://doi.org/10.1111/j.1744-6570.1995.tb01767.x

Kaminske, A. N., Kuepper-Tetzel, C. E., Nebel, C. L., Sumeracki, M. A., & Ryan, S. G. (2020). Transfer: A review for biology and the life sciences. *CBE- Life Sciences Education*, *19*(3). https://doi.org/10.1187/cbe.19-11-0227

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). The advantage of abstract examples in learning math. *Science*, *320*(5875), 454–455. https://doi.org/10.1126/science.1154659

Kant, I. (1958). *Critique of pure reason*, *translated by Norman Kemp Smith*. Macmillan.

Kapur, M. (2010). Productive failure in mathematical problem solving. *Instructional Science*, *38*, 523–550. https://doi.org/10.1007/s11251-009-9093-x

Kapur, M. (2012). Productive failure in learning the concept of variance. *Instructional Science*, *40*, 651–672. https://doi.org/10.1007/s11251-012-9209-6

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. MIT Press.

Kassab, S. E., Al-Shafei, A. I., Salem, A. H., & Otoom, S. (2015). Relationships between the quality of blended learning experience, self-regulated learning, and academic achievement of medical students: A path analysis. *Advances in Medical Education and Practice*, *6*, 27–34. https://doi.org/10.2147/AMEP.S75830

Katayama, A. D., Shambaugh, R. N., & Doctor, T. (2005). Promoting knowledge transfer with electronic note taking. *Teaching of Psychology*, *32*(2), 129–131. https://doi.org/10.1207/s15328023top3202_9

Katzlberger, T. (2005). *Learning by teaching agents* (Publication No. 3217517). [Doctoral dissertation, Vanderbilt University]. ProQuest LLC.

Keane, M. T. (1996). On adaptation in analogy: Tests of pragmatic importance and adaptability in analogical problem solving. *The Quarterly Journal of Experimental Psychology*, 49(4), 1062–1085. https://doi.org/10.1080/713755671

Keane, M. T., Ledgeway, T., & Duff, S. (1994). Constraints on analogical mapping: A comparison of three models. *Cognitive Science*, *18*(3), 387–438. https://doi.org/10.1207/s15516709cog1803_2

Keil, F. C., Smith, W. C., Simons, D. J., & Levin, D. T. (1998). Two dogmas of conceptual empiricism: Implications for hybrid models of the structure of knowledge. *Cognition, 65*(2-3), 103–135. https://doi.org/10.1016/S0010-0277(97)00041-3

Kendeou, P., & Van Den Broek, P. (2005). The effects of readers' misconceptions on comprehension of scientific text. *Journal of Educational Psychology*, *97*(2), 235–245. https://doi.org/10.1037/0022-0663.97.2.235

Kenny, D. A. (2005). Cross-lagged panel design. In B. Everitt & D. Howell (Eds.), *Encyclopedia of statistics in behavioral science*. John Wiley and Sons, Ltd. https://doi.org/10.1002/0470013192.bsa156

Kerslake, D. (1986). *Fractions: Children's strategies and errors. A report of the strategies and errors in secondary mathematics project*. NFER-NELSON Publishing Company.

Kintsch, W. (1970). *Learning*, *memory*, *and conceptual processes*. Wiley.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.

Klahr, D., Triona, L. M., & Williams, C. (2007). Hands on what? The relative effectiveness of physical versus virtual materials in an engineering design project by middle school children. *Journal of Research in Science Teaching*, *44*(1), 183–203. https://doi.org/10.1002/tea.20152

Koca, F. (2016). Motivation to Learn and Teacher-Student Relationship. *Journal of International Education and Leadership*, *6*(2), Article n2.

Koenka, A. C. (2020). Academic motivation theories revisited: An interactive dialog between motivation scholars on recent contributions, underexplored issues, and future directions. *Contemporary Educational Psychology, 61*, Article 101831. https:// doi. org/ 10. 1016/j. cedps ych. 2019. 101831

Kowalski, P., & Taylor, A. F. (2017). Reducing students' misconceptions with refutational teaching: For long-term retention, comprehension matters. *Scholarship of Teaching and Learning in Psychology*, *3*(2), 90–100. https://doi.org/10.1037/stl0000082

Krapp, A. (2000). Interest and human development during adolescence: An educational-psychological approach. In J. Heckhausen (Ed.), *Motivational psychology of human development* (pp. 109–128). Elsevier.

Krapp, A. (2007). An educational-psychological conceptualisation of interest. *International Journal for Educational and Vocational Guidance*, *7*(1), 5–21. https://doi.org/10.1007/s10775-007-9113-9

Krapp, A., & Fink, B. (1992). The development and function of interests during the critical transition from home to preschool. In K. A. Renninger, S. Hidi & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 397–429). Lawrence Erlbaum Associates, Inc.

Krapp, A., Hidi, S., & Renninger, K. A. (1992). Interest, learning, and development. In K. A. Renninger, S. Hidi & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 3–25). Lawrence Erlbaum Associates, Inc.

Kruglanski, A. W. (1978). Endogenous attribution and intrinsic motivation. In Lepper, M., and

Greene, D. (Eds.), *The hidden costs of reward: New perspectives on the psychology of human motivation*. Erlbaum.

Kump, B., Moskaliuk, J., Cress, U., & Kimmerle, J. (2015). Cognitive foundations of organizational learning: Re-introducing the distinction between declarative and non-declarative knowledge. *Frontiers in Psychology*, *6*, Article 1489.

Kurtz, K. J., Miao, C., & Gentner, D. (2001). Learning by analogical bootstrapping. *The Journal of the Learning Sciences*, *10*(4), 417–446. https://doi.org/10.1207/s15327809jls1004new_2

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*, 495–503. https://doi.org/10.1007/s12564-009-9044-x

Laal, M., & Laal, M. (2012). Collaborative learning: what is it? *Procedia-Social and Behavioral Sciences*, *31*, 491–495. https://doi.org/10.1016/j.sbspro.2011.12.092

Lamnina, M., & Chase, C. C. (2021). Uncertain instruction: effects on curiosity, learning, and transfer. *Instructional Science*, *49*, 661–685. https://doi.org/10.1007/s11251-021-09557-2

Lanza, S. T., & Cooper, B. R. (2016). Latent class analysis for developmental research. *Child Development Perspectives*, *10*(1), 59–64. https://doi.org/10.1111/cdep.12163

Lanza, S. T., Tan, X., & Bray, B. C. (2013). Latent class analysis with distal outcomes: A flexible model-based approach. *Structural Equation Modeling*, *20*, 1–20. https://doi.org/10.1080/10705511.2013.742377

Lauermann, F., Meißner, A., & Steinmayr, R. (2020). Relative importance of intelligence and ability self-concept in predicting test performance and school grades in the math and language arts domains. *Educational Psychologist*, *112*(2), 364–383. https://doi.org/10.1037/edu0000377

Lauritzen, P. (2012). *Conceptual and procedural knowledge of mathematical functions* [Doctoral Dissertation, University of Eastern Finland]. Dissertations in Education, Humanities, and Theology No 34.

Lave, J. (1988). *Cognition in practice: Mind*, *mathematics and culture in everyday life*. Cambridge University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge university press.

Lawson, A. E. (2001). Using the learning cycle to teach biology concepts and reasoning patterns. *Journal of Biological Education*, *35*(4), 165–169. https://doi.org/10.1080/00219266.2001.9655772

Lawson, A. E., Alkhoury, S., Benford, R., Clark, B. R., & Falconer, K. A. (2000). What kinds of scientific concepts exist? Concept construction and intellectual development in college biology. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *37*(9), 996–1018.

Leander, K. M. (2001). "This is our freedom bus going home right now": Producing and hybridizing space-time contexts in pedagogical discourse. *Journal of Literacy Research*, *33*(4), 637–679. https://doi.org/10.1080/10862960109548128

Lee, H., Herbert, R. D., & McAuley, J. H. (2019). Mediation analysis. *Jama*, *321*(7), 697–698. https://doi.org/10.1001/jama.2018.21973

LeFevre, J., & Dixon, P. B. (1986). Do written instructions need examples? *Cognition and Instruction*, *3*(1), 1–30. https://doi.org/10.1207/s1532690xci0301_1

LeFevre, J. A., Smith-Chant, B. L., Fast, L., Skwarchuk, S. L., Sargla, E., Arnup, J. S., Penner-Wilger, M., Bisanz, J., & Kamawar, D. (2006). What counts as knowing? The development of conceptual and procedural knowledge of counting from kindergarten through Grade 2. *Journal of Experimental Child Psychology*, *93*(4), 285–303. https://doi.org/10.1016/j.jecp.2005.11.002

Lehman, D. R., Lempert, R., & Nisbett, R. E. (1988). The effects of graduate training on reasoning: Formal discipline and thinking about everyday-life events. *American Psychologist*, *43*(6), 431–442. https://doi.org/10.1037/0003-066x.43.6.431

Lehman, D. R., & Nisbett, R. E. (1990). A longitudinal study of the effects of undergraduate training on reasoning. *Developmental Psychology*, *26*(6), 952–960. https://doi.org/10.1037/0012-1649.26.6.952

Lemos, M. S., & Veríssimo, L. (2014). The relationships between intrinsic motivation, extrinsic motivation, and achievement, along elementary school. *Procedia-Social and Behavioral Sciences*, *112*, 930–938.

Lenz, K., Dreher, A., Holzäpfel, L., & Wittmann, G. (2020). Are conceptual knowledge and procedural knowledge empirically separable? The case of fractions. *British Journal of Educational Psychology*, *90*(3), 809–829. https://doi.org/10.1111/bjep.12333

Lenz, K., & Wittmann, G. (2020). Individual Differences in Conceptual and Procedural Fraction Knowledge: What Makes the Difference and What Does it Look Like?. *International Electronic Journal of Mathematics Education*, *16*(1), Article em0615. https://doi.org/10.29333/iejme/9282

Lepper, M. R., Greene, D. L., & Nisbett, R. E. (1973). Undermining children’s intrinsic interest with extrinsic reward: A test of the „overjustification“ hypothesis. *Journal of Personality and Social Psychology*, *28*(1), 129–137. https://doi.org/10.1037/h0035519

Li, S.-C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., & Baltes, P. B. (2004). Transformations in the couplings among intellectual abilities and constituent cognitive processes across the life span. *Psychological Science*, *15*(3), 155–163.

Lin, Y., McKeachie, W. J., & Kim, Y. S. (2003). College student intrinsic and/or extrinsic motivation and learning. *Learning and Individual Differences*, *13*(3), 251–258. https://doi.org/10.1016/s1041-6080(02)00092-4

Linnenbrink, E. A., & Pintrich, P. R. (2003). The role of self-efficacy beliefs in student engagement and learning in the classroom. *Reading & Writing Quarterly*, *19*(2), 119–137. https://doi.org/10.1080/10573560308223

Little, T. D. (2013). *Longitudinal structural equation modeling*. Guilford Press.

Liu, D., Blickensderfer, E. L., Macchiarella, N. D., & Vincenzi, D. A. (2008). Transfer of training. In D.A. Vincenzi, J.A. Wise, M. Mouloua, & P.A. Hancock (eds), *Human factors in simulation and training*, (pp. 49–60). CRC Press https://doi.org/10.1201/9781420072846

Liu, Y., & Hou, S. (2018). Potential reciprocal relationship between motivation and achievement: A longitudinal study. *School Psychology International*, *39*(1), 38–55. https://doi.org/10.1177/0143034317710574

Lobato, J. (2003). How design experiments can inform a rethinking of transfer and vice versa. *Educational Researcher*, *32*(1), 17–20.

Lobato, J. (2006). Alternative perspectives on the transfer of learning: History, issues, and challenges for future research. *The Journal of the Learning Sciences*, *15*(4), 431–449. https://doi.org/10.1207/s15327809jls1504_1

Lobato, J. (2008). When students don't apply the knowledge you think they have, rethink your assumptions about transfer. In M. P. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and teaching in undergraduate mathematics education* (1st Ed, pp. 289–304). Mathematical Association of America.

Lobato, J. (2012). The actor-oriented transfer perspective and its contributions to educational research and practice. *Educational Psychologist*, *47*(3), 232–247. https://doi.org/10.1080/00461520.2012.693353

Lobato, J., Rhodehamel, B., & Hohensee, C. (2012). "Noticing" as an alternative transfer of learning process. *Journal of the Learning Sciences*, *21*(3), 433–482. https://doi.org/10.1080/10508406.2012.682189

Lobato, J., & Siebert, D. (2002). Quantitative reasoning in a reconceived view of transfer. *The*

Journal of Mathematical Behavior, 21(1), 87–116. https://doi.org/10.1016/S0732-3123(02)00105-0

Locke, E. A., & Schattke, K. (2019). Intrinsic and extrinsic motivation: Time for expansion and clarification. *Motivation Science*, *5*(4), 277–290. https://doi.org/10.1037/mot0000116

Loewenstein, J., Thompson, L., & Gentner, D. (1999). Analogical encoding facilitates knowledge transfer in negotiation. *Psychonomic Bulletin & Review*, *6*(4), 586–597. https://doi.org/10.3758/bf03212967

Longworth, G. (2009). Rationalism and empiricism. In S. Chapman & C. Routledge (Eds.), *Key ideas in linguistics and the philosophy of language* (pp. 67–74). Edinburgh University Press.

López-López, J. A., Page, M. J., Lipsey, M. W., & Higgins, J. P. (2018). Dealing with effect size multiplicity in systematic reviews and meta-analyses. *Research Synthesis Methods*, *9*(3), 336–351. https://doi.org/10.1002/jrsm.1310

Lucas, R. E. (2023). Why the Cross-Lagged Panel Model is Almost Never the Right Choice. *Advances in Methods and Practices in Psychological Science*, *6*(1), 1–22. https://doi.org/10.1177/25152459231158378

Lucas C. G., Bridgers S., Griffiths T. L., Gopnik A. (2014). When children are better (or at least more open-minded) learners than adults: Developmental differences in learning the forms of causal relationships. *Cognition*, 131, 284–299. https://doi.org/10.1016/j.cognition.2013.12.010

Lüdtke, O., & Robitzsch, A. (2022). A comparison of different approaches for estimating cross-lagged effects from a causal inference perspective. *Structural Equation Modeling: A Multidisciplinary Journal*, *29*(6), 888–907. https://doi.org/10.1080/10705511.2022.2065278

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, *58*, 593–614. https://doi.org/10.1146/annurev.psych.58.110405.085542

Marsh, H. W. (2007). *Self-concept theory, measurement and research into practice: The role of self-concept in educational psychology*. British Psychology Society.

Marsh, H. W. (2014). Academic self-concept: Theory, measurement, and research. In H. W. Marsh (Ed.) *Psychological Perspectives on the Self (Vol. 4)* (pp. 71–110), Psychology Press. https://doi.org/10.4324/9781315806976-9

Marsh, H. W., Byrne, B. M., & Yeung, A. S. (1999). Causal ordering of academic self-concept and achievement: Reanalysis of a pioneering study and revised recommendations. *Educational Psychologist*, 34, 155–167. doi:10.1207/s15326985ep3403 2

Marsh, H. W., & Craven, R. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science*, *1*(2), 133–163. https://doi.org/10.1111/j.1745-6916.2006.00010.x

Marsh, H. W., Hau, K.-T., Artelt, C., Baumert, J., & Peschar, J. L. (2006). OECD's brief self-report measure of educational psychology's most useful affective constructs: Cross-cultural, psychometric comparisons across 25 countries [Special Issue]. *International Journal of Testing*, *6*, 311–360.

Marsh, H. W., Hau, K.-T., & Craven, R. (2004). The Big-Fish-Little-Pond Effect Stands Up to Scrutiny. *American Psychologist*, *59*(4), 269–271. https://doi.org/10.1037/0003-066X.59.4.269

Marsh, H. W., Kuyper, H., Seaton, M., Parker, P. D., Morin, A. J. S., Möller, J. et al (2014). Dimensional comparison theory: An extension of the internal/external frame of reference effect on academic self-concept formation. *Contemporary Educational Psychology*, *39*(4), 326–341. https://doi.org/10.1016/j.cedpsych.2014.08.003

Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement:

Relations and causal ordering. *British Journal of Educational Psychology*, *81*(1), 59–77. https://doi.org/10.1348/000709910x503501

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, *76*(2), 397–416. https://doi.org/10.1111/j.1467-8624.2005.00853.x

Marsh, H. W., & Yeung, A. S. (1997). Coursework selection: The effects of academic self-concept and achievement. *American Educational Research Journal*, *34*(4), 691–720. https://doi.org/10.2307/1163354

Marsh, H. W., & Yeung, A. S. (1998). Top-down, bottom-up, and horizontal models: The direction of causality in multidimensional, hierarchical self-concept models. *Journal of Personality and Social Psychology*, *75*(2), 509–527. https://doi.org/10.1037/0022-3514.75.2.509

Marton, F. (2006). Sameness and difference in transfer. *The Journal of the Learning Sciences*, *15*(4), 499–535. https://doi.org/10.1207/s15327809jls1504_3

Matsui, T., Matsui, K., & Ohnishi, R. (1990). Mechanisms underlying math self-efficacy learning of college students. *Journal of Vocational Behavior*, *37*(2), 225–238. https://doi.org/10.1016/0001-8791(90)90042-z

Matthews, P. G., & Fuchs, L. S. (2020). Keys to the gate? Equal sign knowledge at second grade predicts fourth-grade algebra competence. *Child Development*, *91*(1), e14–e28. https://doi.org/10.1111/cdev.13144

Mayer, R. E., & Wittrock, M. C. (1996). Problem-solving transfer. In D. C. Berliner, & R. C. Calfee (Eds.), *Handbook of Educational Psychology* (pp. 47–62). Routledge.

McCarthy, K. S., & McNamara, D. S. (2021). The multidimensional knowledge in text comprehension framework. *Educational Psychologist*, *56*(3), 196–214. https://doi.org/10.1080/00461520.2021.1872379

McCoach, D. B., & Siegle, D. (2001). A comparison of high achievers' and low achievers' attitudes, perceptions, and motivations. *Academic Exchange*, *2*, 71–76.

McCormick, R. (1997). Conceptual and procedural knowledge. *International Journal of Technology and Design Education*, *7*, 141–159. https://doi.org/10.1023/A:1008819912213

McCormick, R. (2009). Learning and teaching: An introduction. In A. T. Jones & M. J. de Vries (Eds.), *International handbook of research and development in technology education* (pp. 363–371). Sense.

McGeoch, J. A. (1942). *The psychology of human learning*. Longmans, Green and Co.

McNamara, D. S., & McDaniel, M. A. (2004). Suppressing irrelevant information: Knowledge activation or inhibition? *Journal of Experimental Psychology: Learning*, *Memory*, *and Cognition*, *30*(2), 465–482. https://doi.org/10.1037/0278-7393.30.2.465

McNeil, N. M., Chesney, D. L., Matthews, P. G., Fyfe, E. R., Petersen, L. A., Dunwiddie, A. E., & Wheeler, M. C. (2012). It pays to be organized: Organizing arithmetic practice around equivalent values facilitates understanding of math equivalence. *Journal of Educational Psychology*, *104*(4), 1109–1121. https://doi.org/10.1037/a0028997

McNeil, N. M., Fyfe, E. R., & Dunwiddie, A. E. (2015). Arithmetic practice can be modified to promote understanding of mathematical equivalence. *Journal of Educational Psychology, 107*(2), 423–436. https://doi.org/10.1037/a0037687

McNeil, N. M., Hornburg, C. B., Brletic-Shipley, H., & Matthews, J. M. (2019). Improving children's understanding of mathematical equivalence via an intervention that goes beyond nontraditional arithmetic practice. *Journal of Educational Psychology*, *111*(6), 1023–1044. https://doi.org/10.1037/edu0000337

McNeil, N. M., Uttal, D. H., Jarvin, L., & Sternberg, R. J. (2009). Should you show me the money? Concrete objects both hurt and help performance on mathematics problems.

Learning and Instruction, *19*(2), 171–184. https://doi.org/10.1016/j.learninstruc.2008.03.005

McNeill, K. L., & Krajcik, J. (2008). Scientific explanations: Characterizing and evaluating the effects of teachers' instructional practices on student learning. *Journal of Research in Science Teaching*, *45(1)*, 53–78. https://doi.org/10.1002/tea.20201

McWeeny, K. H., Young, A. W., Hay, D. C., & Ellis, A. W. (1987). Putting names to faces. *British Journal of Psychology*, *78*(2), 143–149. https://doi.org/10.1111/j.2044-8295.1987.tb02235.x

Meier, C., Frank, C., Gröben, B., & Schack, T. (2020). Verbal instructions and motor learning: How analogy and explicit instructions influence the development of mental representations and tennis serve performance. *Frontiers in Psychology*, *11*, 1–12. https://doi.org/10.3389/fpsyg.2020.00002

Midgley, C. (1993). Motivation and middle level schools. In P. R. Pintrich & M. L. Maehr (Eds.), *Advances in motivation and achievement, Vol. 8: Motivation in the adolescent years* (pp. 219–276). JAI Press.

Möhring, A., Wilhelm, O., & Wilhelm, O. (2018). Knowledge is power for medical assistants: Crystallized and fluid Intelligence as predictors of vocational knowledge. *Frontiers in Psychology*, *9*, Article 28. https://doi.org/10.3389/fpsyg.2018.00028

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. https://doi.org/10.1007/s11409-009-9045-5

Morris, S. C. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods*, *11*(2), 364–386. https://doi.org/10.1177/1094428106291059

Mullins, D., Rummel, N., & Spada, H. (2011). Are two heads always better than one? Differential effects of collaboration on students' computer-supported learning in mathematics. *International Journal of Computer-Supported Collaborative Learning*, *6*, 421–443. https://doi.org/10.1007/s11412-011-9122-z

Munakata, Y., Herd, S. A., Chatham, C. H., Depue, B. E., Banich, M. T., & O'Reilly, R. C. (2011). A unified framework for inhibitory control. Trends in Cognitive Sciences, 15(10), 453–459. https://doi.org/10.1016/j.tics.2011.07.011

Mund, M., & Nestler, S. (2019). Beyond the cross-lagged panel model: Next-generation statistical tools for analyzing interdependencies across the life course. *Advances in Life Course Research*, *41*, Article 100249. https://doi.org/10.1016/j.alcr.2018.10.002

Murphy, K. (2017). What inferences can and cannot be made on the basis of meta-analysis? *Human Resource Management Review*, *27*(1), 193–200. https://doi.org/10.1016/j.hrmr.2015.06.001

Murphy, P. K., & Alexander, P. A. (2000). A motivated exploration of motivation terminology. *Contemporary Educational Psychology*, *25*(1), 3–53. https://doi.org/10.1006/ceps.1999.1019

Muthén, B. O. (2003). Statistical and substantive checking in growth mixture modeling: Comment on Bauer and Curran (2003). *Psychological Methods*, *8*(3), 369–377. https://doi.org/10.1037/1082-989X.8.3.369

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

National Research Council. (2012). *A framework for K-12 science education: Practices*, *crosscutting concepts*, *and core ideas*. National Academies Press.

NCTM. (2014). *Principles to actions: Ensuring mathematical success for all*. National Council of Teachers of Mathematics, Inc.

Neely, J. H. (1989). Experimental dissociations and the episodic/semantic memory distinction.

In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 229–270). Lawrence Erlbaum Associates, Inc.

Nelson, W., Reyna, V. F., Fagerlin, A., Lipkus, I., & Peters, E. (2008). Clinical implications of numeracy: theory and practice. *Annals of Behavioral Medicine*, *35*(3), 261–274. https://doi.org/10.1007/s12160-008-9037-8

Nesher, P. (1986). Learning mathematics: A cognitive perspective. *American Psychologist*, *41*(10), 1114–1122. https://doi.org/10.1037/0003-066x.41.10.1114

Neta, R., & Pritchard, D. (Eds.) (2009). *Arguing About Knowledge*. Routledge.

Newsom, J. T. (2015). Cross-lagged panel analysis. In S. Krauss Whitbourne (Ed.), *The Encyclopedia of Adulthood and Aging*. John Wiley and Sons, Inc. https://doi.org/10.1002/9781118521373.wbeaa250

Ngu, B. H., Mit, E., Shahbodin, F., & Tuovinen, J. (2009). Chemistry problem solving instruction: a comparison of three computer-based formats for learning from hierarchical network problem representations. *Instructional Science*, *37*, 21–42. https://doi.org/10.1007/s11251-008-9072-7

Ngu, B. H., & Phan, H. P. (2016). Comparing balance and inverse methods on learning conceptual and procedural knowledge in equation solving: a cognitive load perspective. *Pedagogies: An International Journal*, *11*(1), 63–83. https://doi.org/10.1080/1554480X.2015.1047836

Nisbett, R. E., Evans, M. I., Jepson, C., & Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review*, *90*(4), 339–363. https://doi.org/10.1037/0033-295x.90.4.339

Noche, J. R., Baloloy, E. T., De Luna, Y. V. V., Salcedo, F. C. P., & Abellon, E. C. (2019). The impact of supplemental self-paced instruction on conceptual understanding and procedural skill of adolescents. *The Mathematics Educator 18*(1-2), 75–99.

Nokes, T. J. (2009). Mechanisms of knowledge transfer. *Thinking & Reasoning*, *15*(1), 1–36. https://doi.org/10.1080/13546780802490186

Nokes-Malach, T. J., & Mestre, J. P. (2013). Toward a model of transfer as sense-making. *Educational Psychologist*, *48*(3), 184–207. https://doi.org/10.1080/00461520.2013.807556

Nonaka, I., & Takeuchi, H. (1995). *The Knowledge-Creating Company: How Japanese companies create dynamics of innovation*. Oxford University Press.

Nussenbaum, K., & Hartley, C. A. (2019). Reinforcement learning across development: What insights can we draw from a decade of research? Developmental Cognitive Neuroscience, 40, Article 100733. https://doi.org/10.1016/j.dcn.2019.100733

Nutbrown, C. (2006). *Threads of thinking: Young children learning and the role of early education*. Sage.

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. https://doi.org/10.1080/10705510701575396

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

Obiagu, A. N., Mezieobi, D. I., Aroh, P. N., & Akubue, F. N. (2020). The effect of cooperative concept mapping on misconceptions, knowledge achievement, and transfer of learning in peace education. *The Social Studies*, *111*(1), 18–38. https://doi.org/10.1080/00377996.2019.1652139

Ohlsson, S. (1994). Declarative and procedural knowledge. In T. Husen & T. Neville-Postlethwaite (Eds.), *The international encyclopedia of education* (Vol. 3 , 2nd ed., pp. 1432–1434). Pergamon Press.

Ohlsson, S. (1996). Learning from performance errors. *Psychological Review*, *103*(2), 241–262. https://doi.org/10.1037/0033-295x.103.2.241

Ohlsson, S., Ernst, A., & Rees, E. (1992). The cognitive complexity of learning and doing arithmetic. *Journal for Research in Mathematics Education*, *23*(5), 441–467. https://doi.org/10.2307/749565

Ormrod, J. E. (2012). *Human Learning* (6th ed). Pearson Education (US).

Ormrod, J. E., Ormrod, R., Wagner, E. D., & McCallin, R. C. (1988). Reconceptualizing map learning. *American Journal of Psychology*, *101*(3), 425–433. https://doi.org/10.2307/1423088

Orth, U., Clark, D. A., Donnellan, M. B., & Robins, R. W. (2021). Testing prospective effects in longitudinal research: Comparing seven competing cross-lagged models. *Journal of Personality and Social Psychology*, *120*(4), 1013–1034. http://dx.doi.org/10.1037/pspp0000358

Orth, U., Meier, L. L., Bühler, J. L., Dapp, L. C., Krauss, S., Messerli, D., & Robins, R. W. (2022). Effect size guidelines for cross-lagged effects. *Psychological Methods.* Advance online publication. https://doi.org/10.1037/met0000499

Osborne, J. W., & Jones, B. D. (2011). Identification with academics and motivation to achieve in school: How the structure of the self influences academic outcomes. *Educational Psychology Review*, *23*, 131–158. https://doi.org/10.1007/s10648-011-9151-1

Osgood, C. E. (1949). The similarity paradox in human learning: A resolution. *Psychological Review*, *56*(3), 132–143. https://doi.org/10.1037/h0057488

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *International Journal of Surgery*, *88*, Article 105906.

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., & Mulrow, C. D. (2021). PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ*, *372*(160), 1–36. https://doi.org/10.1136/bmj.n160

Pai, H., Sears, D. A., & Maeda, Y. (2015). Effects of small-group learning on transfer: A meta-analysis. *Educational Psychology Review*, *27*(1), 79–102. https://doi.org/10.1007/s10648-014-9260-8

Pan, S. C., & Rickard, T. C. (2018). Transfer of test-enhanced learning: Meta-analytic review and synthesis. *Psychological Bulletin*, *144*(7), 710–756. https://doi.org/10.1037/bul0000151

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

Paul, J. M., Gray, S. A., Butterworth, B. L., & Reeve, R. A. (2019). Reading and math tests differentially predict number transcoding and number fact speed longitudinally: A random intercept cross-lagged panel approach. *Journal of Educational Psychology*, *111*(2), 299–313. https://doi.org/10.1037/edu0000287

Pazzani, M. J. (1991). Influence of prior knowledge on concept acquisition: Experimental and computational results. *Journal of Experimental Psychology: Learning*, *Memory and Cognition*, *17*(3), 416–432. https://doi.org/10.1037/0278-7393.17.3.416

Peiffer, H., Ellwart, T., & Preckel, F. (2020). Ability self-concept and self-efficacy in higher education: An empirical differentiation based on their factorial structure. *PLoS One*, *15*(7), Article e0234604.

Peled, I., & Segalis, B. (2005). It's not too late to conceptualize: Constructing a generalized subtraction schema by abstracting and connecting procedures. *Mathematical Thinking and Learning*, *7*(3), 207–230. https://doi.org/10.1207/s15327833mtl0703_2

Pellas, N., Fotaris, P., Kazanidis, I., & Wells, D. (2019). Augmenting the learning experience in primary and secondary school education: A systematic review of recent trends in augmented reality game-based learning. *Virtual Reality*, *23*(4), 329–346. https://doi.org/10.1007/s10055-018-0347-2

Perkins, D. N. (1992). *Smart schools: Better thinking and learning for every child.* The Free Press.

Perkins, D. N., & Salomon, G. (1992). Transfer of learning. In T. Husén & T. N. Postlethwaite (Eds.), *International Encyclopedia of Education*, *2* (pp. 6452–6457). Elsevier Science.

Perkins, D. N., & Salomon, G. (2012). Knowledge to go: A motivational and dispositional view of transfer. *Educational Psychologist*, *47*(3), 248–258. https://doi.org/10.1080/00461520.2012.693354

Phye, G. (1994). Learning and remembering: The basis for personal knowledge construction. In G. Phye (Ed.), *Handbook of Academic Learning* (pp. 47–64). Academic Press.

Pintrich, P. R., & Zusho, A. (2002). The development of academic self-regulation: The role of cognitive and motivational factors. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 249–284), Academic.

Ploger, D., & Hecht, S. A. (2009). Enhancing children's conceptual understanding of mathematics through chartworld software. *Journal of Research in Childhood Education*, *23*(3), 267–277. https://doi.org/10.1080/02568540909594660

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. https://doi.org/10.1007/s11145-010-9260-0

Pritchard, D. (2013). *What is this thing called knowledge?* Routledge.

Pustejovsky, J., & Pustejovsky, M. J. (2020). *Package 'clubSandwich'*. CRAN.

Quillian, M. R. (1968). Semantic memory. In M. Minsky (Ed.), *Semantic information processing* (pp. 227–270). MIT Press.

R Core Team (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Ratelle, C. F., Guay, F., Vallerand, R. J., Larose, S., & Senecal, C. (2007). Autonomous, controlled, and amotivated types of academic motivation: A person-oriented analysis. *Journal of Educational Psychology*, *99*(4), 734–746. https://doi.org/10.1037/0022-0663.99.4.734

Rau, M. A., Aleven, V., & Rummel, N. (2015). Successful learning with multiple graphical representations and self-explanation prompts. *Journal of Educational Psychology*, *107*(1), 30–46. https://doi.org/10.1037/a0037211

Rawson, K. A., Thomas, R. C., & Jacoby, L. L. (2015). The power of examples: Illustrative examples enhance conceptual learning of declarative concepts. *Educational Psychology Review*, *27*, 483–504. https://doi.org/10.1007/s10648-014-9273-3

Reed, S. K. (2012). Learning by mapping across situations. *The Journal of the Learning Sciences*, *21*(3), 353–398. https://doi.org/10.1080/10508406.2011.607007

Reed, S. K., Ernst, G. W., & Banerji, R. B. (1974). The role of analogy in transfer between similar problem states. *Cognitive Psychology*, *6*(3), 436–450. https://doi.org/10.1016/0010-0285(74)90020-6

Reeves, L. M., & Weisberg, R. H. (1994). The role of content and abstract information in analogical transfer. *Psychological Bulletin*, *115*(3), 381–400. https://doi.org/10.1037/0033-2909.115.3.381

Reigeluth, C. M. (1983). Instructional design: What is it and why is it. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: An overview of their current status* (pp. 3–36). Routledge.

Reigeluth, C. M. (1999). What is instructional-design theory and how is it changing? In C. M.

Reigeluth (Ed.), *Instructional-design theories and models: A new paradigm of instructional theory* (Vol. II, pp. 5–29). Lawrence Erlbaum Associates

Reigeluth, C. M., & An, Y. (2020). *Merging the instructional design process with learner-centered theory: The holistic 4D model*. Routledge.

Reigeluth, C. M., & Carr-Chellman, A. A. (2009). Understanding Instructional Theory, C. M. Reigeluth & A. A. Carr-Chellman (Eds.), *Instructional-design theories and models, volume III: Building a common knowledge base (Vol. 3)* (pp. 3–26). Routledge.

Reigeluth, C. M., & Merrill, M. D. (1979). Classes of instructional variables. *Educational Technology*, *19*(3), 5–24.

Reiser, R. A. (2018). A history of instructional design and technology. In R.A. Reiser & J.V. Dempsey (Eds.), *Trends and issues in instructional design and technology* (4th ed, pp. 17–34). Pearson Education

Renkl, A. (2009). Wissenserwerb. In E. Wild & Möller, J. (Eds). *Pädagogische Psychologie* (pp. 3–24), Springer-Lehrbuch, Springer.

Renninger, K. A., & Hidi, S. (2002). Student interest and achievement: Developmental issues raised by a case study. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 173–195)., New York: Academic.

Renninger, K. A., & Hidi, S. (2011). Revisiting the conceptualization, measurement, and generation of interest. *Educational Psychologist*, *46*(3), 168–184. https://doi.org/10.1080/00461520.2011.587723

Resnick, L. B., & Ford, W. W. (1981). *The psychology of mathematics for instruction*. Routledge.

Resnick, L. B., & Omanson, S. F. (1987). Leaming to understand arithmetic. In Glaser, R. (Ed.), *Advances in instructional psychology (Vol. 3)* (pp. 41–95). Erlbaum.

Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin*, *135*(6), 943–973. https://doi.org/10.1037/a0017327

Reyna, V. F., Nelson, W. L., Han, P. K., & Pignone, M. P. (2015). Decision making and cancer. *American Psychologist*, *70*(2), 105–118. https://doi.org/10.1037/a0036834

Rips, L., & Medin, D. (2005). Concepts and categories: Memory, meaning and metaphysics. In K. J. Holyoak & B. G. Morrison (Eds.), *Cambridge handbook of thinking and reasoning*. Cambridge University Press.

Ritchie, S. J., & Bates, T. C. (2013). Enduring links from childhood mathematics and reading achievement to adult socioeconomic status. *Psychological Science*, *24*(7), 1301–1308. https://doi.org/10.1177/0956797612466268

Rittle-Johnson, B. (2006). Promoting transfer: Effects of self-explanation and direct instruction. *Child Development*, *77*(1), 1–15. https://doi.org/10.1111/j.1467-8624.2006.00852.x

Rittle-Johnson, B. (2017). Developing mathematics knowledge. *Child Development Perspectives*, *11*(3), 184–190. https://doi.org/10.1111/cdep.12229

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other?. *Journal of Educational Psychology*, *91*(1), 175–189. https://doi.org/10.1037/0022-0663.91.1.175

Rittle-Johnson, B., Fyfe, E. R., & Loehr, A. M. (2016). Improving conceptual and procedural knowledge: The impact of instructional content within a mathematics lesson. *British Journal of Educational Psychology*, *86*(4), 576–591. https://doi.org/10.1111/bjep.12124

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. https://doi.org/10.1348/000709908X398106

Rittle-Johnson, B., Loehr, A. M., & Durkin, K. (2017). Promoting self-explanation to improve mathematics learning: A meta-analysis and instructional design principles. *ZDM*, *49*, 599–611. https://doi.org/10.1007/s11858-017-0834-z

Rittle-Johnson, B., & Schneider, M. (2014). Developing conceptual and procedural knowledge of mathematics. In R. C. Kadosh & A. Dowker (Eds.), *Oxford Handbook of numerical cognition*, (pp. 1118–1134). Oxford University Press.

Rittle-Johnson, B., Schneider, M., & Star, J. R. (2015). Not a one-way street: Bidirectional relations between procedural and conceptual knowledge of mathematics. *Educational Psychology Review*, *27*(4), 587–597. https://doi.org/10.1007/s10648-015-9302-x

Rittle-Johnson, B., & Siegler, R. S. (1998). The relation between conceptual and procedural knowledge in learning mathematics: A review. In C. Donlan (Ed.), *The development of mathematical skills*, (pp. 75–110). Psychology Press.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, *93*(2), 346–362. https://doi.org/10.1037/0022-0663.93.2.346

Robinson, K. M., & Dubé, A. K. (2009). A microgenetic study of the multiplication and division inversion concept. *Canadian Journal of Experimental Psychology*, *63*(3), 193–200. https://doi.org/10.1037/a0013908

Rohrer, D., & Pashler, H. (2010). Recent research on human learning challenges conventional instructional strategies. *Educational Researcher*, *39*(5), 406–412. https://doi.org/10.3102/0013189x10374770

Roorda, G., Vos, P., & Goedhart, M. (2015). An actor-oriented transfer perspective on high school students' development of the use of procedures to solve problems on rate of change. *International Journal of Science and Mathematics Education*, *13*(4), 863–889. https://doi.org/10.1007/s10763-013-9501-1

Rosch, E., Simpson, C., & Miller, R. S. (1976). Structural bases of typicality effects. *Journal* *of Experimental Psychology: Human Perception and Performance, 2*(4), 491–502. https://doi.org/10.1037/0096-1523.2.4.491

Roscoe, R. D., Gutierrez, P. J., Wylie, R., & Chi, M. T. H. (2014). Evaluating lesson design and implementation within the ICAP framework. *Boulder*, *CO: International Society of the Learning Sciences*. https://repository.isls.org//handle/1/1223

Rosel, J., & Plewis, I. (2008). Longitudinal data analysis with structural equations. *Methodology*, *4*(1), 37–50. https://doi.org/10.1027/1614-2241.4.1.37

Rosenberg, M. (1979). *Conceiving the Self*. Basic Books.

Rotgans, J. I., & Schmidt, H. G. (2017). The role of interest in learning: knowledge acquisition at the intersection of situational and individual interest. In P. A. O'Keefe & J. M. Harackiewicz (Eds.), *The science of interest* (pp. 69–93). Springer. https://doi.org/10.1007/978-3-319-55509-6_4

Roth, W. M., & Jornet, A. (2013). Situated cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(5), 463–478. https://doi.org/10.1002/wcs.1242

Royer, J. M. (1979). Theories of the transfer of learning. *Educational Psychologist*, *14*(1), 53–69. https://doi.org/10.1080/00461527909529207

Russell, B. (1972). *A history of western philosophy*. Simon and Schuster.

Ryan, R. M., & Deci, E. L. (2000a). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Ryan, R. M., & Deci, E. L. (2000b). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037//0003-066X.55.1.68

Ryan, R. M., & Deci, E. L. (2002). Overview of self-determination theory: An organismic-dialectical perspective. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 3–33). University of Rochester Press.

Ryan, R. M., Deci, E. L., Vansteenkiste, M., & Soenens, B. (2021). Building a science of motivated persons: Self-determination theory's empirical approach to human experience and the regulation of behavior. *Motivation Science*, *7*(2), 97–110. https://doi.org/10.1037/mot0000194

Sackett, P. R., Harris, M. E., & Orr, J. (1986). On seeking moderator variables in the meta-analysis of correlational data: A Monte Carlo investigation of statistical power and resistance to type I error. *Journal of Applied Psychology*, *71*(2), 302–310. https://doi.org/10.1037/0021-9010.71.2.302

Sacks, H. S., Berrier, J., Reitman, D., Ancona-Berk, V., & Chalmers, T. C. (1987). Meta-analyses of randomized controlled trials. *The New England Journal of Medicine*, *316*(8), 450–455. https://doi.org/10.1056/nejm198702193160806

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119–144. https://doi.org/10.1007/bf00117714

Salomon, G., & Perkins, D. L. (1989). Rocky roads to transfer: Rethinking mechanism of a neglected phenomenon. *Educational Psychologist*, *24*(2), 113–142. https://doi.org/10.1207/s15326985ep2402_1

Sansone, C., & Smith, J. L. (2000). Interest and self-regulation: The relation between having to and wanting to. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic and Extrinsic Motivation: The Search for Optimal Motivation and Performance* (pp. 341–372), Academic.

Sayer, A. (2010). Reductionism in social science. In R. L. Lee (Ed.), *Questioning Nineteenth-Century Assumptions about Knowledge, II: Reductionism* (pp. 7–39). State University of New York Press.

Scammacca, N., Roberts, G., Vaughn, S., Edmonds, M., Wexler, J., Reutebuch, C. K., & Torgesen, J. K. (2007). Interventions for adolescent struggling readers: A meta-analysis with implications for practice. *Center on Instruction*.

Scammacca, N. K., Roberts, G., Vaughn, S., & Stuebing, K. K. (2015). A meta-analysis of interventions for struggling readers in grades 4–12: 1980–2011. *Journal of Learning Disabilities, 48*, 369–390. https://doi.org/10.1177/0022219413504995

Schalk, L., Saalbach, H., & Stern, E. (2016). Approaches to foster transfer of formal principles: which route to take? *PloS one*, *11*(2), Article e0148787. https://doi.org/10.1371/journal.pone.0148787

Scherer, R., Siddiq, F., & Sánchez Viveros, B. (2019). The cognitive benefits of learning computer programming: A meta-analysis of transfer effects. *Journal of Educational Psychology*, *111*(5), 764–792. https://doi.org/10.1037/edu0000314

Schiefele, U. (1990). The influence of topic interest, prior knowledge, and cognitive capabilities on text comprehension. In J. M. Pieters, K. Breuer, & P. R. J. Simons (Eds.), *Learning environments* (pp. 323–337). Springer.

Schiefele, U. (1991). Interest, learning, and motivation. *Educational Psychologist*, *26*(3), 299–323. http://dx.doi.org/10.1207/s15326985ep2603&4_5

Schiefele, U. (1992). Topic interest and levels of text comprehension. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 151–182). Erlbaum.

Schiefele, U., & Krapp, A. (1991, April). *The effects of topic interest and cognitive characteristics on different indicators of free recall of expository text* [Paper presentation]. Annual Meeting of the American Educational Research Association, Chicago.

Schiefele, U., Krapp, A., & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and developement* (pp. 183–212). Psychology Press.

Schiepe-Tiska, A., Rönnebeck, S., Schöps, K., Neumann, K., Schmidtner, S., Parchmann, I.,

& Prenzel, M. (2015). Naturwissenschaftliche Kompetenz in PISA 2015 – Ergebnisse des internationalen Vergleichs mit einem modifizierten Testansatz. In K. Reiss, C. Sälzer, A.Schiepe-Tiska, E. Klieme & O. Köller (Eds.), *PISA 2015. Eine Studie Zwischen Kontinuität Und Innovation* (pp. 45–98). Waxmann.

Schiepe-Tiska, A., Simm, I., & Schmidtner, S. (2015). Motivationale Orientierungen, Selbstbilder und Berufserwartungen in den Naturwissenschaften in PISA 2015. In K. Reiss, C. Sälzer, A.Schiepe-Tiska, E. Klieme & O. Köller (Eds.), *PISA 2015. Eine Studie Zwischen Kontinuität Und Innovation* (pp. 99–132). Waxmann.

Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods*, *1*(2), 199–223. https://doi.org/10.1037/1082-989x.1.2.199

Schmidt, F. L., & Hunter, J. E. (2015). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd ed.). Sage Publications, Inc.

Schneider, M., Grabner, R. H., & Paetsch, J. (2009). Mental number line, number line estimation, and mathematical achievement: their interrelations in grades 5 and 6. *Journal of Educational Psychology*, *101*(2), 359–372. https://doi.org/10.1037/a0013840

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

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*, *47*(6), 1525–1538. https://doi.org/10.1037/a0024997

Schneider, M., & Stern, E. (2009). The inverse relation of addition and subtraction: A knowledge integration perspective. *Mathematical Thinking and Learning*, *11*(1-2), 92–101. https://doi.org/10.1080/10986060802584012

Schneider, M., & Stern, E. (2010). The developmental relations between conceptual and procedural knowledge: A multimethod approach. *Developmental Psychology*, *46*(1), 178–192. https://doi.org/10.1037/a0016701

Schneider, S., Nebel, S., Pradel, S., & Rey, G. D. (2015). Introducing the familiarity mechanism: A unified explanatory approach for the personalization effect and the examination of youth slang in multimedia learning. *Computers in Human Behavior*, *43*, 129–138. https://doi.org/10.1016/j.chb.2014.10.052

Schneider, W. (1993). Domain-specific knowledge and memory performance in children. *Educational Psychology Review*. https://doi.org/10.1007/bf01323047

Schofield, H. (2013). The philosophy of education (RLE Edu K): an introduction. *Routledge*.

Schunk, D. H. (1985). Self-efficacy and classroom learning. *Psychology in the Schools*, *22*(2), 208–223. https://doi.org/10.1002/1520-6807(198504)22:2

Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist*, *26*(3-4), 207–231. https://doi.org/10.1080/00461520.1991.9653133

Schunk, D. H., Meece, J. L., & Pintrich, P. R. (2014). *Motivation in education: Theory, research, and applications* (4th ed.). Pearson.

Schunk, D. H., & Pajares, F. (2002). The development of academic self-efficacy. In A. Wigfield & J. S. Eccles (Eds.), *Development of Achievement Motivation* (pp. 15–31), Academic Press.

Schwartz, D. A., Varma, S., & Martin, L. (2008). Dynamic transfer and innovation. In S. Vosniadou (Ed.), *International Handbook of Research on Conceptual Change* (pp. 479–507). Routledge.

Schwartz, D. K., & Martin, T. (2004). Inventing to prepare for future Learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction*, *22*(2), 129–184. https://doi.org/10.1207/s1532690xci2202_1

Schwartz, D. L., Bransford, J. D., & Sears, D. (2005). Efficiency and innovation in transfer. In J. Mestre (Ed.), *Transfer of Learning: Research and Perspectives* (pp. 1–51). Information Age Publishing Inc.

Schwartz, D. L., Sears, D., & Chang, J. (2007). Reconsidering prior knowledge. In M. C. Lovett & P. Shah (Eds.), *Thinking with Data* (pp. 319–344). Psychology Press.

Schwarz, B. B., Neuman, Y., & Biezuner, S. (2000). Two wrongs may make a right . . . If they argue together! *Cognition and Instruction*, *18*(4), 461–494. https://doi.org/10.1207/s1532690xci1804_2

Schwarzer, G., Carpenter, J. R., & Rücker, G. (2015). *Meta-Analysis with R*. Springer. https://doi.org/10.1007/978-3-319-21416-0

Seels, B., & Glasgow, Z. (1998). *Making Instructional Design Decisions* (2nd ed.). Prentice Hall.

Seiver E., Gopnik A., Goodman N. D. (2013). Did she jump because she was the big sister or because the trampoline was safe? Causal inference and the development of social attribution. *Child Development*, 84, 443–454. https://doi.org/10.1111/j.1467-8624.2012.01865.x

Selig, J. P., & Little, T. D. (2012). Autoregressive and cross-lagged panel analysis for longitudinal data. In B. Laursen, T. D. Little, & N. A. Card (Eds.), *Handbook of Developmental Research Methods* (pp. 265–278). The Guilford Press.

Sénéchal, M., LeFevre, J. A., Smith-Chant, B. L., & Colton, K. V. (2001). On refining theoretical models of emergent literacy the role of empirical evidence. *Journal of School Psychology*, *39*(5), 439–460. https://doi.org/10.1016/S0022-4405(01)00081-4

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

Shastri, L. (1985). *Evidential Reasoning in Semantic Networks: A Formal Theory and Its Parallel Implementation*. (Publication No. 8528562) [Doctoral Dissertation, University of Rochester]. ProQuest LLC.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, *46*(3), 407–441. https://doi.org/10.3102/00346543046003407

Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(2), 179–187. https://doi.org/10.1037/0278-7393.5.2.179

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*(2), 127–190. https://doi.org/10.1037/0033-295x.84.2.127

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology*, *70*(1), 747–770. https://doi.org/10.1146/annurev-psych-010418-102803

Siegler, R. S., & Crowley, K. (1994). Constraints on learning in nonprivileged domains. *Cognitive Psychology, 27*(2), 194–226. https://doi.org/10.1006/cogp.1994.1016

Siegler, R. S., & Stern, E. (1998). Conscious and unconscious strategy discoveries: A microgenetic analysis. *Journal of Experimental Psychology*, *127*(4), 377–397. https://doi.org/10.1037/0096-3445.127.4.377

Silver, E. A. (1994). On mathematical problem posing. *For the Learning of Mathematics*, *14*(1), 19–28.

Simon, Y. R. (2020). *An Introduction to Metaphysics of Knowledge*. Fordham University Press.

Simonsmeier, B. A., Flaig, M., Deiglmayr, A., Schalk, L., & Schneider, M. (2022). Domain-specific prior knowledge and learning: A meta-analysis. *Educational Psychologist*, *57*(1), 31–54. https://doi.org/10.1080/00461520.2021.1939700

Singley, M. K., & Anderson, J. R. (1989). *The Transfer of Cognitive Skill (Cognitive Science Series*, *9)* (1. Ed.). Harvard University Press.

Skinner, B. (1953). *Science And Human Behavior*. Macmillan.

Sloutsky, V. M., Kaminski, J. A., & Heckler, A. F. (2005). The advantage of simple symbols for learning and transfer. *Psychonomic Bulletin & Review*, *12*(3), 508–513. https://doi.org/10.3758/BF03193796

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. https://doi.org/10.1207/s15327809jls0302_1

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, *13*, 290–312. http://doi.org/10.2307/270723

Song, F., Parekh, S., Hooper, L., Loke, Y., Ryder, J., Sutton, A., Hing, C., Kwok, C. S., Pang, C., & Harvey, I. (2010). Dissemination and publication of research findings: an updated review of related biases. *Health Technology Assessment*, *14*(8). https://doi.org/10.3310/hta14080

Sowa, J. F. (1984). *Conceptual Structures: Information Processing in Mind and Machine*. Addison Wesley

Sowa, J. F. (2014). Part I Issues in Semantic Networks. In J. F. Sowa (Ed.), *Principles of Semantic Networks: Explorations in the Representation of Knowledge* (pp. 1–11). Morgan Kaufmann.

Spelke, E. S. (2004). Core knowledge. In N. Kanwisher & J. Duncan (Eds.), *Attention and performance*, *vol.20: Functional neuroimaging of visual cognition*. Oxford University Press.

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, *10*(1), 89–96. https://doi.org/10.1111/j.1467-7687.2007.00569.x

Spurk, D., Hirschi, A., Wang, M., Valero, D., & Kauffeld, S. (2020). Latent profile analysis: A review and "how to" guide of its application within vocational behavior research. *Journal of Vocational Behavior*, *120*, Article 103445. https://doi.org/10.1016/j.jvb.2020.103445

Srisawasdi, N., & Panjaburee, P. (2019). Implementation of game-transformed inquiry-based learning to promote the understanding of and motivation to learn chemistry. *Journal of Science Education and Technology*, *28*, 152–164. https://doi.org/10.1007/s10956-018-9754-0

Star, J. R. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education*, *36*(5), 404–411. https://doi.org/10.2307/30034943

Star, J. R. (2007). Foregrounding procedural knowledge. *Journal for Research in Mathematics Education*, *38*(2), 132–135.

Star, J. R., & Stylianides, G. J. (2013). Procedural and conceptual knowledge: Exploring the gap between knowledge type and knowledge quality. *Canadian Journal of Science, Mathematics and Technology Education*, *13*, 169–181.

Sternberg, R. J. (1981). Nothing fails like success: The search for an intelligent paradigm for studying intelligence. *Journal of Educational Psychology, 73*(2), 142–155. https://doi.org/10.1037/0022-0663.73.2.142

Sterne, J. A., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H.-Y., Corbett, M. S., Eldridge, S. M., Emberson, J. R., Hernán, M. A., Hopewell, S., Hróbjartsson, A., Junqueira, D. R., Jüni, P., Kirkham, J. J., Lasserson, T., Li, T., …, Higgins, J. P. T. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*, *366*, Article 14898. https://doi.org/10.1136/bmj.l4898

Stratton, D. H. (2020). Types of instructional strategies and their effect on Preparation for

Future Learning in differentiation. *International Journal of Educational Research*, *104*, Article 101691. https://doi.org/10.1016/j.ijer.2020.101691

Subedi, B. S. (2004). Emerging trends of research on transfer of learning. *International Education Journal*, *5*(4), 591–599.

Suchman, L. A. (1987). *Plans and Situated Actions*. Cambridge University Press.

Suggate, S. P. (2016). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities*, *49*(1), 77–96. https://doi.org/10.1177/0022219414528540

Taasoobshirazi, G., & Carr, M. (2008). A review and critique of context-based physics instruction and assessment. *Educational Research Review*, *3*(2), 155–167. https://doi.org/10.1016/j.edurev.2008.01.002

Taatgen, N., & Anderson, J. R. (2002). Why do children learn to say “Broke”? A model of learning the past tense without feedback. *Cognition*, *86*(2), 123–155. https://doi.org/10.1016/s0010-0277(02)00176-2

Tannenbaum, S. I., & Yukl, G. (1992). Training and development in work organizations. *Annual Review of Psychology*, *43*(1), 399–441. https://doi.org/10.1146/annurev.ps.43.020192.002151

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. https://doi.org/10.1007/s40865-016-0026-5

Taylor, G., Jungert, T., Mageau, G. A., Schattke, K., Dedic, H., Rosenfield, S., & Koestner, R. (2014). A self-determination theory approach to predicting school achievement over time: the unique role of intrinsic motivation. *Contemporary Educational Psychology*, *39*(4), 342–358. https://doi.org/10.1016/j.cedpsych.2014.08.002

Ten Berge, T., & Van Hezewijk, R. (1999). Procedural and declarative knowledge: An evolutionary perspective. *Theory & Psychology*, *9*(5), 605–624.

Thomas, R., Anderson, L., Getahun, L., & Cooke, B. (1992). Teaching for transfer of learning. *National Center for Research in Vocational Education*, *Berkeley*, *CA*.

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. https://doi.org/10.1207/S15328023TOP3002_02

Thompson, S. G., & Pocock, S. J. (1991). Can meta-analyses be trusted? *The Lancet*, *338*(8775), 1127–1130. https://doi.org/10.1016/0140-6736(91)91975-z

Thompson-Schill, S. L., Ramscar, M., & Chrysikou, E. G. (2009). Cognition without control: When a little frontal lobe goes a long way. *Current Directions in Psychological Science, 18*(5), 259–263. https://10.1111/j.1467-8721.2009.01648.x

Thorndike, E. L. (1924). Mental discipline in high school studies. *Journal of Educational Psychology*, *15*(2), 83–98. https://doi.org/10.1037/h0071035

Tizard, B., & Hughes, M. (2008). *Young children learning*. John Wiley & Sons.

Tobias, S. (1982). When do instructional methods. *Educational Researcher, 11*(4), 4–9.

Tobias, S. (1994). Interest, prior knowledge, and learning *Review of Educational Research*, *64*(1), 37–54. https://doi.org/10.3102/00346543064001037

Toll, S. W., & Van Luit, J. E. H. (2014). Effects of remedial numeracy instruction throughout kindergarten starting at different ages: Evidence from a large-scale longitudinal study. *Learning and Instruction*, *33*, 39–49. https://doi.org/10.1016/j.learninstruc.2014.03.003

Trafton, J. G., & Trickett, S. B. (2001). Note-taking for self-explanation and problem solving. *Human-Computer Interaction*, *16*(1), 1–38. https://doi.org/10.1207/s15327051hci1601_1

Treagust, D. F. (2014). General instructional methods and strategies. In N. G. Lederman & S.

K. Abell (Eds.), *Handbook of research on science education* (pp. 373–391). Routledge.

Tulving, E. (1993). What is episodic memory? *Current Directions in Psychological Science*, *2*(3), 67–70. https://doi.org/10.1111/1467-8721.ep10770899

Urhahne, D., & Wijnia, L. (2023). Theories of motivation in education: An integrative framework. *Educational Psychology Review*, *35*(2), 45.

Usami, S., Murayama, K., & Hamaker, E. L. (2019). A unified framework of longitudinal models to examine reciprocal relations. *Psychological Methods*, *24*(5), 637–657. http://dx.doi.org/10.1037/met0000210

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, *39*(2), 111–133. https://doi.org/10.1207/s15326985ep3902_3

Valenza, E., Leo, I., Gava, L., & Simion, F. (2006). Perceptual completion in new born Human infants. *Child Development*, *77*(6), 1810–1821. https://doi.org/10.1111/j.1467-8624.2006.00975.x

Vallerand, R. J., Fortier, M. S., & Guay, F. (1997). Self-determination and persistence in a real-life setting: Toward a motivational model of high school dropout. Journal of Personality and Social Psychology, 72, 1161–1176.

van Merriënboer, J. J. G. (2012). Variability of practice. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 3389–3390). Springer. https://doi.org/10.1007/978-1-4419-1428-6_415

van Merriënboer, J. J. G., & Kester, L. (2014). The four-component instructional design model: Multimedia principles in environments for complex learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 104–148). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.007

van Merriënboer, J. J. G., & Kirschner, P. A. (2017). *Ten steps to complex learning: A systematic approach to four-component instructional design.* Routledge.

Van Meter, P., & Garner, J. K. (2005). The promise and practice of learner-generated drawing: Literature review and synthesis. *Educational Psychology Review*, *17*(4), 285–325. https://doi.org/10.1007/s10648-005-8136-3

VanderWeele, T. J. (2016). Mediation analysis: a practitioner's guide. *Annual Review of Public Health*, *37*, 17–32. https://doi.org/10.1146/annurev-publhealth-032315-021402

Vanhove, A. J., & Harms, P. D. (2015). Reconciling the two disciplines of organisational science: A comparison of findings from lab and field research. *Applied Psychology*, *64*(4), 637–673. https://doi.org/10.1111/apps.12046

VanScoy, A. (2019). Conceptual and procedural knowledge: A framework for analyzing point-of-need information literacy instruction. *Communications in Information Literacy*, *13*(2), 164–180. https://doi.org/10.15760/comminfolit.2019.13.2.3

Velada, R., Caetano, A., Michel, J. W., Lyons, B. D., & Kavanagh, M. J. (2007). The effects of training design, individual characteristics and work environment on transfer of training. *International Journal of Training and Development*, *11*(4), 282–294. https://doi.org/10.1111/j.1468-2419.2007.00286.x

Vermunt, J. K. (2002). Latent class analysis of complex sample survey data: Application to dietary data: Comment. *Journal of the American Statistical Association*, *97*(459), 736–737. https://doi.org/10.1177/0095798420930932

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3). 1–48. https://doi.org/10.18637/jss.v036.i03

Viechtbauer, W., & Viechtbauer, M. W. (2015). *Package 'metafor'. The Comprehensive R Archive Network. Package 'metafor'.* http://cran.r-project.org/web/packages/metafor/metafor.pdf

Viechtbauer, W., & Cheung M. W. L. (2010) Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, *1*(2), 112–125. https://doi.org/10.1002/jrsm.11

Vosniadou, S., & Ortony, A. (1989). Similarity and analogical reasoning: a synthesis. In S.

Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 1–18). Cambridge University Press. https://doi.org/10.1017/cbo9780511529863.002

Voss, J. F., Blais, J., Means, M. L., Greene, T. C., & Ahwesh, E. R. (1986). Informal reasoning and subject matter knowledge in the solving of economics problems by naive and novice individuals. *Cognition and Instruction*, *3*(3), 269–302. https://doi.org/10.1207/s1532690xci0303_7

Vu, T., Magis-Weinberg, L., Jansen, B. R., van Atteveldt, N., Janssen, T. W., Lee, N. C., van der Maas, H. L. J., Raijmakers, M. E. J., Sachisthal, M. S. M., & Meeter, M. (2022). Motivation-achievement cycles in learning: A literature review and research agenda. *Educational Psychology Review*, *34*(1), 39–71.

Wagensveld, B., Segers, E., Kleemans, T., & Verhoeven, L. (2015). Child predictors of learning to control variables via instruction or self-discovery. *Instructional Science*, *43*, 365–379. https://doi.org/10.1007/s11251-014-9334-5

Walker, E., Hernandez, A. F., & Kattan, M. W. (2008). Meta-analysis: Its strengths and limitations. *Cleveland Clinic Journal of Medicine*, *75*(6), 431–439. https://doi.org/10.3949/ccjm.75.6.431

Walter, S. D., Irwig, L., & Glasziou, P. P. (1999). Meta-analysis of diagnostic tests with imperfect reference standards. *Journal of Clinical Epidemiology*, *52*(10), 943–951. https://doi.org/10.1016/S0895-4356(99)00086-4

Wasiu, O. I., & Abiola, O. A. (2019). Enhancing the conceptual, procedural and flexible procedural knowledge of pre-service mathematics teachers in algebra. *Journal of Research and Advances in Mathematics Education*, *4*(2), 66–78.

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

Weaver, J. C., Chastain, R. J., DeCaro, D. A., & DeCaro, M. S. (2018). Reverse the routine: Problem solving before instruction improves conceptual knowledge in undergraduate physics. *Contemporary Educational Psychology*, *52*, 36–47. https://doi.org/10.1016/j.cedpsych.2017.12.003

Webb, N. M. (1989). Peer interaction, problem-solving, and cognition: Multidisciplinary perspectives. *International Journal of Educational Research*, *13*(1), 1. https://doi.org/10.1016/0883-0355(89)90009-8

Weinfurt, K. P., Castel, L. D., Li, Y., Sulmasy, D. P., Balshem, A. M., Benson III, A. B., Burnett, C. B., Gaskin, D. J., Marshall, J. L., Slater, E. F., Schulman, K. A., Meropol, N. J. (2003). The correlation between patient characteristics and expectations of benefit from Phase I clinical trials. *Cancer*, *98*(1), 166–175. https://doi.org/10.1002/cncr.11483

Weinstein, C. E. (1987). Fostering learning autonomy through the use of learning strategies. *Journal of Reading*, *30*(7), 590–595.

Weller, B. E., Bowen, N. K., & Faubert, S. J. (2020). Latent class analysis: a guide to best practice. *Journal of Black Psychology*, *46*(4), 287–311. https://doi.org/10.1177/0095798420930932

Wendling, B. J., Mather, N., & Schrank, F. A. (2009). Woodcock-Johnson III tests of cognitive abilities. In J. A. Naglieri & S. Goldstein (Eds.), *Practitioner's Guide to assessing intelligence and achievement*. (pp. 191–229). John Wiley & Sons, Inc.

Wenger, E. (1998). *Communities of Practice: Learning*, *Meaning*, *and Identity*. Cambridge University Press.

Wenger, E., McDermott, R. A., & Snyder, W. (2002). *Cultivating Communities of Practice*. Harvard Business School Press.

Whitehead, A. N. (1967). *The aims of education and other essays*. Macmillan.

Wiernik, B. M., & Dahlke, J. A. (2020). Obtaining unbiased results in meta-analysis: The importance of correcting for statistical artifacts. *Advances in Methods and Practices in Psychological Science*, *3*(1), 94–123. https://doi.org/10.1177/2515245919885611

Wigfield, A., & Karpathian, M. (1991). Who am I and what can I do? Children's self-concepts and motivation in achievement situations. *Educational Psychologist*, *26*(3-4), 233–261. https://doi.org/10.1080/00461520.1991.9653134

Wiggins, B. L., Eddy, S. L., Grunspan, D. Z., & Crowe, A. J. (2017). The ICAP Active Learning framework predicts the learning gains observed in intensely active classroom experiences. *AERA Open*, *3*(2). 1–14. https://doi.org/10.1177/2332858417708567

Williams, P. B., & Carnine, D. W. (1981). Relationship between range of examples and of instructions and attention in concept attainment. *The Journal of Educational Research*, *74*(3), 144–148. https://doi.org/10.1080/00220671.1981.10885300

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625–636. https://doi.org/10.3758/bf03196322

Wilson, S. J., Polanin, J. R., & Lipsey, M. W. (2016). Fitting meta-analytic structural equation models with complex datasets. *Research Synthesis Methods*, *7*(2), 121–139. https://doi.org/10.1002/jrsm.1199

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: a meta-analysis of educational feedback research. *Frontiers in Psychology*, *10*, Article 3087. https://doi.org/10.3389/fpsyg.2019.03087

Witherby, A. E., & Carpenter, S. K. (2021). The rich-get-richer effect: Prior knowledge predicts new learning of domain-relevant information. *Journal of Experimental Psychology: Learning*, *Memory*, *and Cognition*. https://doi.org/10.1037/xlm0000996

Woodworth, R. S., & Thorndike, E. H. (1901). The influence of improvement in one mental function upon the efficiency of other functions. (I). *Psychological Review*, *8*(3), 247–261. https://doi.org/10.1037/h0074898

Wu, H., Guo, Y., Yang, Y., Zhao, L., & Guo, C. (2021). A meta-analysis of the longitudinal relationship between academic self-concept and academic achievement. *Educational Psychology Review, 33*, 1749-1778. https://doi.org/10.1007/s10648-021-09600-1

Yang, H. T., & Wang, K. H. (2014). A teaching model for scaffolding 4th grade students' scientific explanation writing. *Research in Science Education*, *44*, 531–548. https://doi.org/10.1007/s11165-013-9392-8

Ye, A., Resnick, I., Hansen, N., Rodrigues, J., Rinne, L., & Jordan, N. C. (2016). Pathways to fraction learning: Numerical abilities mediate the relation between early cognitive competencies and later fraction knowledge. *Journal of Experimental Child Psychology*, *152*, 242–263. https://doi.org/10.1016/j.jecp.2016.08.001

Yin, Z. G., Zhang, J. B., Kan, S. L., & Wang, X. G. (2012). Diagnostic accuracy of imaging modalities for suspected scaphoid fractures: meta-analysis combined with latent class analysis. *The Journal of Bone and Joint Surgery*. *British Volume*, *94*(8), 1077–1085. https://doi.org/10.1302/0301-620x.94b8.28998

Yuan, Y., & MacKinnon, D. P. (2009). Bayesian mediation analysis. *Psychological Methods*, *14*, 301–322. http://doi.org/10.1037/a0016972

Zhang, T., Chen, A., Yli-Piipari, S., Loflin, J., Wells, S., Schweighardt, R., Moennich, K., Hong, D., & Ennis, C. D. (2016). Prior knowledge determines interest in learning in physical education: A structural growth model perspective. *Learning and Individual Differences*, *51*, 132–140. https://doi.org/10.1016/j.lindif.2016.08.039

Zhao, X., Lynch Jr, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, *37*(2), 197–206. https://doi.org/10.1086/651257

Zoupidis, A., Pnevmatikos, D., Spyrtou, A., & Kariotoglou, P. (2016). The impact of procedural and epistemological knowledge on conceptual understanding: the case of density and floating-sinking phenomena. *Instructional Science*, *44*(4), 315-334. https://doi.org/10.1007/s11251-016-9375-z

Appendix

Appendix A: Included Studies (Study 1)

Achmetli, K., Schukajlow, S., & Rakoczy, K. (2019). Multiple solutions for real-world problems, experience of competence and students' procedural and conceptual knowledge. *International Journal of Science and Mathematics Education*, *17*(8), 1605–1625. https://doi.org/10.1007/s10763-018-9936-5

Bailey, D. H., Hansen, N., & Jordan, N. C. (2017). The codevelopment of children's fraction arithmetic skill and fraction magnitude understanding. *Journal of Educational Psychology*, *109*(4), 509–519. https://doi.org/10.1037/edu0000152

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

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

Ching (2016). *The importance of additive reasoning in children's mathematical achievement: A longitudinal study*. (https://ora.ox.ac.uk/objects/uuid:65ec7289-d41d-4634-8a99-9eda2210c4a4). [Doctoral Dissertation, University of Oxford]. Boby Ching.

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*, Article 775. https://doi.org/10.3389/fpsyg.2016.00775

Cowan, R., Donlan, C., Shepherd, D.-L., Cole-Fletcher, R., Saxton, M., & Hurry, J. (2011). Basic calculation proficiency and mathematics achievement in elementary school children. *Journal of Educational Psychology*, *103*, 786–803. https://doi.org/10.1037/a0024556.

Fyfe, E. R., DeCaro, M. S., & Rittle-Johnson, B. (2014). An alternative time for telling: When conceptual instruction prior to problem solving improves mathematical knowledge. *British Journal of Educational Psychology*, *84*(3), 502–519. https://doi.org/10.1111/bjep.12035

Geary, D. C., Vanmarle, K., Chu, F. W., Hoard, M. K., & Nugent, L. (2019). Predicting age of becoming a cardinal principle knower. *Journal of Educational Psychology*, *111*(2), 256–267. https://psycnet.apa.org/doi/10.1037/edu0000277

Gunderson, E. A., & Hildebrand, L. (2021). Relations among spatial skills, number line estimation, and exact and approximate calculation in young children. *Journal of Experimental Child Psychology*, *212*, 105251. https://doi.org/10.1016/j.jecp.2021.105251

Hansen, N., Jordan, N. C., Fernandez, E., Siegler, R. S., Fuchs, L., Gersten, R., & Micklos, D. (2015). General and math-specific predictors of sixth-graders' knowledge of fractions. *Cognitive Development*, *35*, 34–49.

Hansen, N., Jordan, N. C., & Rodrigues, J. (2017). Identifying learning difficulties with fractions: A longitudinal study of student growth from third through sixth grade. *Contemporary Educational Psychology*, *50*, 45–59. https://doi.org/10.1016/j.cogdev.2015.02.001

Hecht, S. A., & Vagi, K. J. (2010). Sources of group and individual differences in emerging fraction skills. *Journal of Educational Psychology*, *102*(4), 843–859. https://doi.org/10.1037/a0019824

Heemsoth, T., & Heinze, A. (2016). Secondary school students learning from reflections on the rationale behind self-made errors: A field experiment. *The Journal of Experimental Education*, *84*(1), 98–118. https://doi.org/10.1080/00220973.2014.963215

Jõgi, A. L., & Kikas, E. (2016). Calculation and word problem-solving skills in primary grades-Impact of cognitive abilities and longitudinal interrelations with task-persistent behaviour. *British Journal of Educational Psychology*, *86*(2), 165–181. https://doi.org/10.1111/bjep.12096

Jordan, C. L., Hansen, N., Fuchs, L., Siegler, R., Micklos, D., & Gersten, R. (2013). Developmental predictors of conceptual and procedural knowledge of fractions. *Journal of Experimental Child Psychology*, *116*(1), 45–58. https://doi.org/10.1016/j.jecp.2013.02.001

Koponen, T., Aunola, K., Ahonen, T., & Nurmi, J. E. (2007). Cognitive predictors of single-digit and procedural calculation skills and their covariation with reading skill. *Journal of Experimental Child Psychology*, *97*(3), 220–241. https://doi.org/10.1016/j.jecp.2007.03.001

Koponen, T., Salmi, P., Torppa, M., Eklund, K., Aro, T., Aro, M., ... & Nurmi, J. E. (2016). Counting and rapid naming predict the fluency of arithmetic and reading skills. *Contemporary Educational Psychology*, *44*, 83–94. https://doi.org/10.1016/j.cedpsych.2016.02.004

Laski, E. V., Schiffman, J., Shen, C., & Vasilyeva, M. (2016). Kindergartners' base-10 knowledge predicts arithmetic accuracy concurrently and longitudinally. *Learning and Individual Differences*, *50*, 234–239. https://doi.org/10.1016/j.lindif.2016.08.004

LeFevre, J. A., Berrigan, L., Vendetti, C., Kamawar, D., Bisanz, J., Skwarchuk, S. L., & Smith-Chant, B. L. (2013). The role of executive attention in the acquisition of mathematical skills for children in Grades 2 through 4. *Journal of Experimental Child Psychology*, *114*(2), 243–261. https://doi.org/10.1016/j.jecp.2012.10.005

Ngu, B. H., & Phan, H. P. (2016). Comparing balance and inverse methods on learning conceptual and procedural knowledge in equation solving: a cognitive load perspective. *Pedagogies: An International Journal*, *11*(1), 63–83. https://doi.org/10.1080/1554480X.2015.1047836

Noche, J. R., Baloloy, E. T., De Luna, Y. V. V., Salcedo, F. C. P., & Abellon, E. C. (2019). The impact of supplemental self-paced instruction on conceptual understanding and procedural skill of adolescents. *The Mathematics Educator 18*(1-2), 75–99.

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

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other?. *Journal of Educational Psychology*, *91*(1), 175–189. https://doi.org/10.1037/0022-0663.91.1.175

Rittle-Johnson, B., Fyfe, E. R., Hofer, K. G., & Farran, D. C. (2017). Early math trajectories: Low-income children's mathematics knowledge from ages 4 to 11. *Child Development*, *88*(5), 1727–1742. https://doi.org/10.1111/cdev.12662

Rittle-Johnson, B., Fyfe, E. R., & Loehr, A. M. (2016). Improving conceptual and procedural knowledge: The impact of instructional content within a mathematics lesson. *British Journal of Educational Psychology*, *86*(4), 576–591. https://doi.org/10.1111/bjep.12124

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. https://doi.org/10.1348/000709908X398106

Rittle-Johnson, B., & Star, J. R. (2009). Compared with what? The effects of different comparisons on conceptual knowledge and procedural flexibility for equation solving. *Journal of Educational Psychology*, *101*(3), 529–544. https://doi.org/10.1037/a0014224

Rittle-Johnson, B., Star, J. R., & Durkin, K. (2011). Developing procedural flexibility: Are novices prepared to learn from comparing procedures? *British Journal of Educational Psychology*, *83(3)*,436–455. https://doi.org/10.1111/j.2044-8279.2011.02037.x

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*, *47*(6), 1525–1538. https://doi.org/10.1037/a0024997

Schneider, M., & Stern, E. (2010). The developmental relations between conceptual and procedural knowledge: A multimethod approach. *Developmental Psychology*, *46*(1), 178–192. https://doi.org/10.1037/a0016701

Star, J. R., Pollack, C., Durkin, K., Rittle-Johnson, B., Lynch, K., Newton, K., & Gogolen, C. (2015). Learning from comparison in algebra. *Contemporary Educational Psychology*, *40*, 41–54. https://doi.org/10.1016/j.cedpsych.2014.05.005

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

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. https://doi.org/10.1111/cdev.12218

Wasiu, O. I., & Abiola, O. A. (2019). Enhancing the conceptual, procedural and flexible procedural knowledge of pre-service mathematics teachers in algebra. *Journal of Research and Advances in Mathematics Education*, *4*(2), 66–78.

Ye, A., Resnick, I., Hansen, N., Rodrigues, J., Rinne, L., & Jordan, N. C. (2016). Pathways to fraction learning: Numerical abilities mediate the relation between early cognitive competencies and later fraction knowledge. *Journal of Experimental Child Psychology*, *152*, 242–263. https://doi.org/10.1016/j.jecp.2016.08.001

Appendix B: Coded Moderator Variables (Study 1)

Table 17

Overview of Subgroup Analyses of Moderator Levels (Study 1)

Moderator level	*j*	*k*	Path coefficient	95% CI	τ^2
Overall model					
$r^+_{C1\text{-}P1}$	30	47	.398	[.354; .443]	.005
$\beta_{C1\text{-}C2}$	25	44	.402	[.346; .458]	.015
$\beta_{C1\text{-}P2}$	27	42	.270	[.221; .318]	.013
$\beta_{P1\text{-}C2}$	25	43	.220	[.172; .268]	.007
$\beta_{P1\text{-}P2}$	25	44	.336	[.291; .380]	.007
$r^+_{\text{C2-P2}}$	22	41	.243	[.113; .354]	.129
Moderator: school level					
Kindergarten / Preschool					
$r^+_{C1\text{-}P1}$	8	9	.467	[.362; .571][a]	.001
$\beta_{C1\text{-}C2}$	5	6	.171	[.078; .264][a]	.002
$\beta_{C1\text{-}P2}$	6	7	.123	[.010; .236][a]	.000
$\beta_{P1\text{-}C2}$	5	6	.312	[.203; .421][a]	.002
$\beta_{P1\text{-}P2}$	6	7	.412	[.324; .571][a]	.005
$r^+_{\text{C2-P2}}$	3	4	.137	[.025; .248][a]	.003
Primary school					
$r^+_{C1\text{-}P1}$	13	17	.410	[.351; .470]	.001
$\beta_{C1\text{-}C2}$	12	18	.421	[.311; .532]	.009
$\beta_{C1\text{-}P2}$	11	17	.330	[.257; .403]	.024
$\beta_{P1\text{-}C2}$	11	15	.186	[.120; .252]	.005
$\beta_{P1\text{-}P2}$	10	14	.272	[.212; .331]	.004
$r^+_{\text{C2-P2}}$	9	15	.346	[.228; .465]	.069
Secondary school					
$r^+_{C1\text{-}P1}$	9	18	.386	[.317; .455]	.012
$\beta_{C1\text{-}C2}$	8	17	.387	[.294; .480]	.010
$\beta_{C1\text{-}P2}$	9	18	.275	[.196; .354]	.016
$\beta_{P1\text{-}C2}$	9	18	.226	[.150; .303]	.008
$\beta_{P1\text{-}P2}$	10	19	.350	[.267; .433]	.009
$r^+_{\text{C2-P2}}$	9	18	.349	[.179; .520]	.137
Moderator: Subdomain					
Whole-number operations					
$r^+_{C1\text{-}P1}$	11	14	.435	[.349; .520]	.000

Table 17 (continued)

Moderator level	j	k	Path coefficient	95% CI	τ^2
$\beta_{C1\text{-}C2}$	6	9	.266	[.131; .402]	.018
$\beta_{C1\text{-}P2}$	10	13	.158	[.064; .253]	.013
$\beta_{P1\text{-}C2}$	6	9	.303	[.196; .410]	.003
$\beta_{P1\text{-}P2}$	10	13	.404	[.342; .466]	.010
$r^+_{C2\text{-}P2}$	5	8	.199	[.024; .374]	.054
Rational numbers					
$r^+_{C1\text{-}P1}$	7	11	.407	-[b]	.017
$\beta_{C1\text{-}C2}$	7	11	.448	-[b]	.013
$\beta_{C1\text{-}P2}$	7	11	.221	-[b]	.058
$\beta_{P1\text{-}C2}$	7	11	.072	-[b]	.008
$\beta_{P1\text{-}P2}$	7	11	.349	-[b]	.013
$r^+_{C2\text{-}P2}$	7	11	.300	-[b]	.021
Algebra					
$r^+_{C1\text{-}P1}$	7	15	.327	[.259; .395]	.000
$\beta_{C1\text{-}C2}$	8	18	.360	[.224; .496]	.009
$\beta_{C1\text{-}P2}$	8	18	.268	[.198; .339]	.054
$\beta_{P1\text{-}C2}$	8	16	.270	[.181; .359]	.000
$\beta_{P1\text{-}P2}$	8	16	.302	[.256; .347]	.011
$r^+_{C2\text{-}P2}$	9	19	.381	[.318; .443]	.019
Moderator: Intervention					
Studies with intervention					
$r^+_{C1\text{-}P1}$	13	25	.364	[.305; .423]	.012
$\beta_{C1\text{-}C2}$	14	26	.413	[.323; .502]	.010
$\beta_{C1\text{-}P2}$	15	27	.280	[.215; .346]	.026
$\beta_{P1\text{-}C2}$	14	25	.212	[.149; .274]	.006
$\beta_{P1\text{-}P2}$	15	26	.337	[.267; .423]	.005
$r^+_{C2\text{-}P2}$	15	27	.319	[.218; .420]	.075
Studies without intervention					
$r^+_{C1\text{-}P1}$	19	22	.425	-[b]	.001
$\beta_{C1\text{-}C2}$	15	18	.398	-[b]	.012
$\beta_{C1\text{-}P2}$	16	19	.267	-[b]	.012
$\beta_{P1\text{-}C2}$	14	17	.224	-[b]	.006
$\beta_{P1\text{-}P2}$	15	18	.332	-[b]	.012
$r^+_{C2\text{-}P2}$	11	14	.159	-[b]	.091
Moderator: Broad Task Type					
Implicit					
$r^+_{C1\text{-}P1}$	23	32	.406	[.352; .460]	.005

Table 17 (continued)

Moderator level	j	k	Path coefficient	95% CI	τ^2
$\beta_{C1\text{-}C2}$	18	29	.357	[.295; .420]	.017
$\beta_{C1\text{-}P2}$	20	31	.269	[.208; .331]	.007
$\beta_{P1\text{-}C2}$	17	26	.204	[.154; .255]	.006
$\beta_{P1\text{-}P2}$	19	28	.325	[.270; .381]	.003
$r^{+}_{C2\text{-}P2}$	14	25	.406	[.207; .404]	.072
Explicit					
$r^{+}_{C1\text{-}P1}$	3	7	.291	- [b]	.001
$\beta_{C1\text{-}C2}$	3	7	.310	- [b]	.000
$\beta_{C1\text{-}P2}$	3	7	.323	- [b]	.005
$\beta_{P1\text{-}C2}$	3	7	.244	- [b]	.000
$\beta_{P1\text{-}P2}$	3	7	.308	- [b]	.000
$r^{+}_{C2\text{-}P2}$	3	7	.240	- [b]	.010
Both					
$r^{+}_{C1\text{-}P1}$	3	6	.453	- [b]	.001
$\beta_{C1\text{-}C2}$	4	9	.590	- [b]	.019
$\beta_{C1\text{-}P2}$	4	9	.279	- [b]	.027
$\beta_{P1\text{-}C2}$	3	6	.212	- [b]	.003
$\beta_{P1\text{-}P2}$	3	6	.265	- [b]	.027
$r^{+}_{C2\text{-}P2}$	4	9	.284	- [b]	.048
Moderator: Specific Task Type					
Translate Quantities					
$r^{+}_{C1\text{-}P1}$	18	24	.459	[.395; .524]	.002
$\beta_{C1\text{-}C2}$	13	19	.404	[.333; .474]	.019
$\beta_{C1\text{-}P2}$	14	20	.270	[.198; .342]	.004
$\beta_{P1\text{-}C2}$	13	19	.177	[.100; .254]	.005
$\beta_{P1\text{-}P2}$	14	20	.303	[.243; .364]	.009
$r^{+}_{C2\text{-}P2}$	9	15	.247	[.131; .364]	.058
Compare Quantities					
$r^{+}_{C1\text{-}P1}$	7	9	.365	[.250; .479]	.001
$\beta_{C1\text{-}C2}$	7	9	.294	[.187; .400]	.002
$\beta_{C1\text{-}P2}$	6	7	.253	[.161; .344]	.010
$\beta_{P1\text{-}C2}$	6	7	.223	[.101; .345]	.004
$\beta_{P1\text{-}P2}$	5	6	.267	[.178; .354]	.010
$r^{+}_{C2\text{-}P2}$	5	6	.277	[.169; .386]	.012
Evaluate Unfamiliar Procedures					
$r^{+}_{C1\text{-}P1}$	2	4	.301	- [b]	.000

Table 17 (continued)

Moderator level	j	k	Path coefficient	95% CI	τ^2
$\beta_{C1\text{-}C2}$	3	7	.241	- [b]	.000
$\beta_{C1\text{-}P2}$	3	7	.238	- [b]	.000
$\beta_{P1\text{-}C2}$	2	4	.156	- [b]	.000
$\beta_{P1\text{-}P2}$	2	4	.202	- [b]	.000
$r^{+}_{C2\text{-}P2}$	3	7	.323	- [b]	.000
Explain Judgments					
$r^{+}_{C1\text{-}P1}$	3	7	.291	- [b]	.001
$\beta_{C1\text{-}C2}$	3	7	.310	- [b]	.000
$\beta_{C1\text{-}P2}$	3	7	.323	- [b]	.004
$\beta_{P1\text{-}C2}$	3	7	.244	- [b]	.000
$\beta_{P1\text{-}P2}$	3	7	.308	- [b]	.000
$r^{+}_{C2\text{-}P2}$	3	7	.240	- [b]	.010
Several					
$r^{+}_{C1\text{-}P1}$	7	13	.383	[.302; .463]	.017
$\beta_{C1\text{-}C2}$	8	16	.481	[.372; .591]	.012
$\beta_{C1\text{-}P2}$	7	13	.324	[.240; .407]	.020
$\beta_{P1\text{-}C2}$	7	15	.211	[.128; .294]	.004
$\beta_{P1\text{-}P2}$	6	12	.343	[.233; .463]	.004
$r^{+}_{C2\text{-}P2}$	7	15	.236	[.129; .342]	.042

Note. Overview of included categorical moderator levels with their respective results in subgroup analyses; j - Number of studies, k - Number of independent correlation matrices, 95% CI - 95% confidence interval, τ^2 - Heterogeneity index.

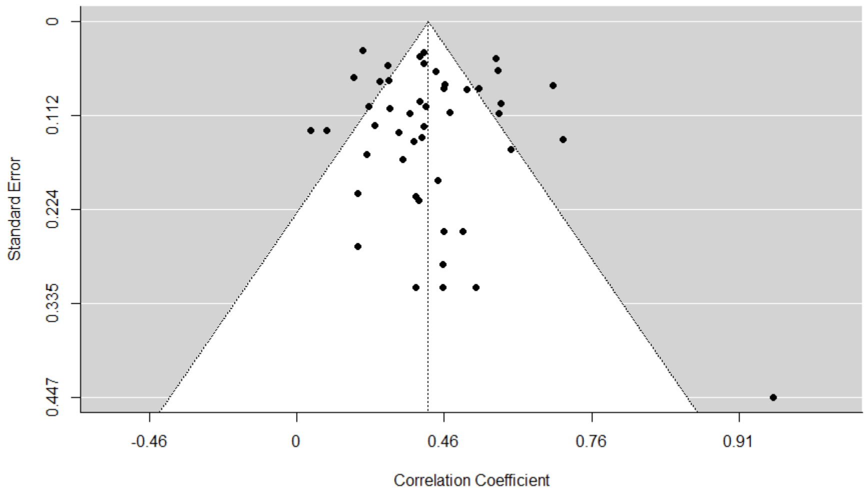
[a] = Confidence interval estimated by setting variances without SEs at 0.

[b] = Estimation of confidence interval not possible.

Appendix C: Funnel Plots of Correlations (Study 1)

Figure 35

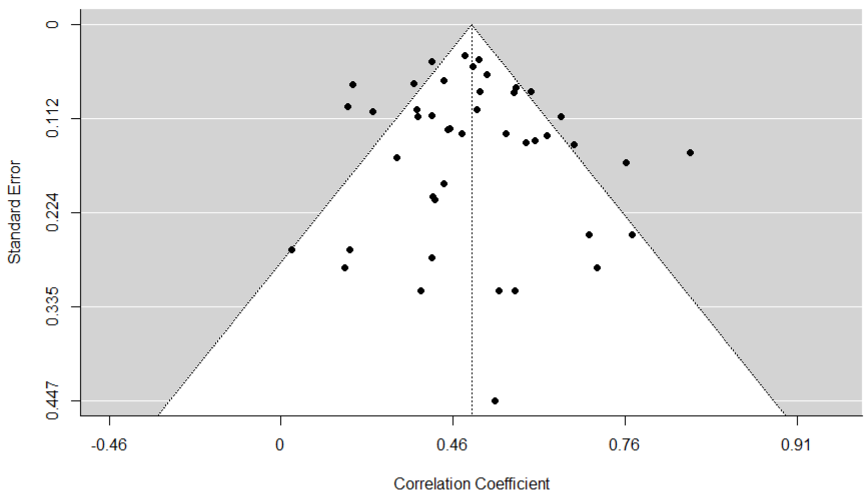
Funnel Plot of the Bivariate Correlations Between Conceptual and Procedural Knowledge at T1



Note. Funnel Plot of the included aggregated bivariate correlations on sample level.

Figure 36

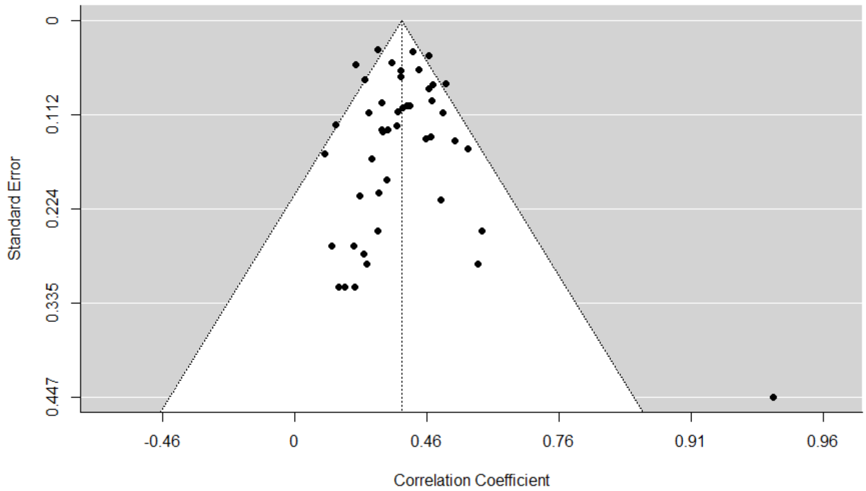
Funnel Plot of the Bivariate Correlations Between Conceptual Knowledge at T1 and T2



Note. Funnel Plot of the included aggregated bivariate correlations on sample level.

Figure 37

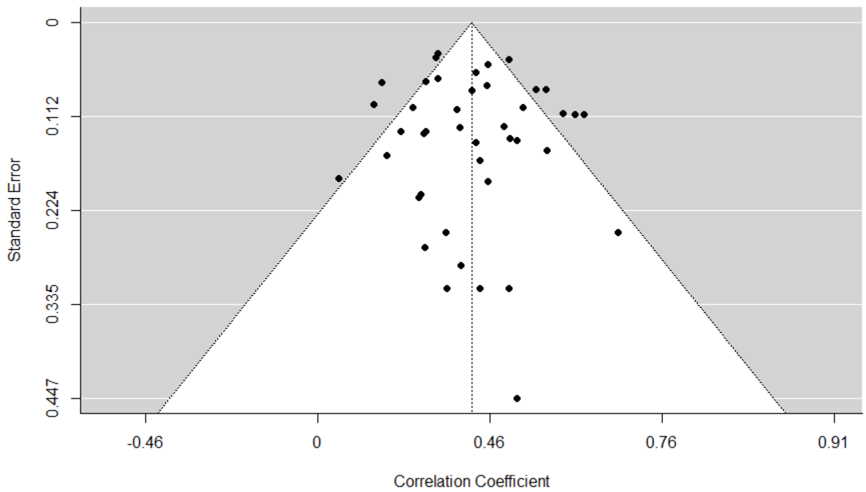
Funnel Plot of the Bivariate Correlations Between Conceptual Knowledge at T1 and Procedural Knowledge at T2



Note. Funnel Plot of the included aggregated bivariate correlations on sample level.

Figure 38

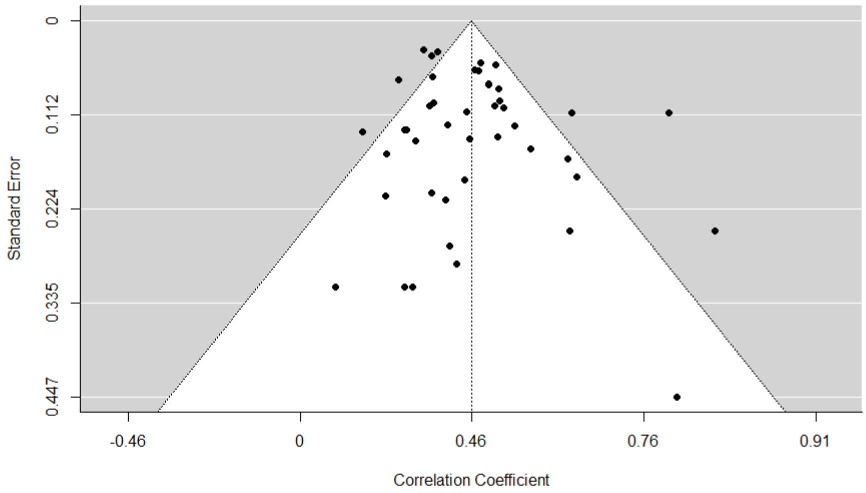
Funnel Plot of the Bivariate Correlations Between Procedural Knowledge at T1 and Conceptual Knowledge at T2



Note. Funnel Plot of the included aggregated bivariate correlations on sample level.

Figure 39

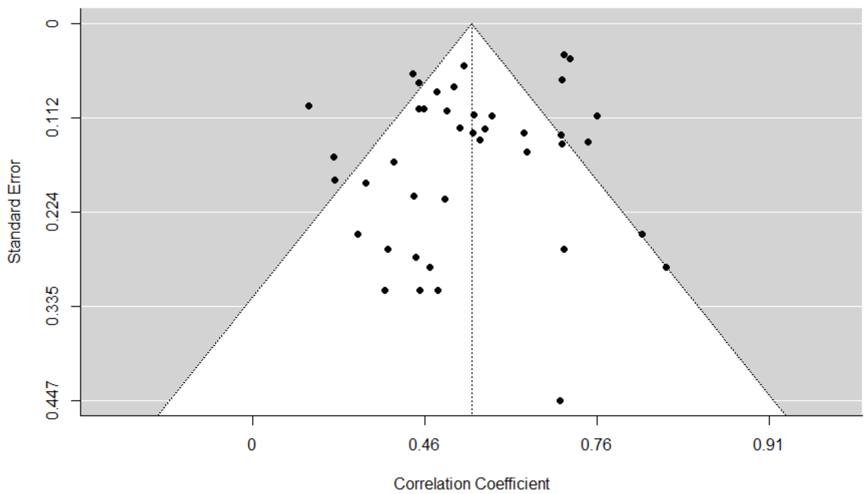
Funnel Plot of the Bivariate Correlations Between Procedural Knowledge at T1 and T2



Note. Funnel Plot of the included aggregated bivariate correlations on sample level.

Figure 40

Funnel Plot of the Bivariate Correlations Between Conceptual and Procedural Knowledge at T2

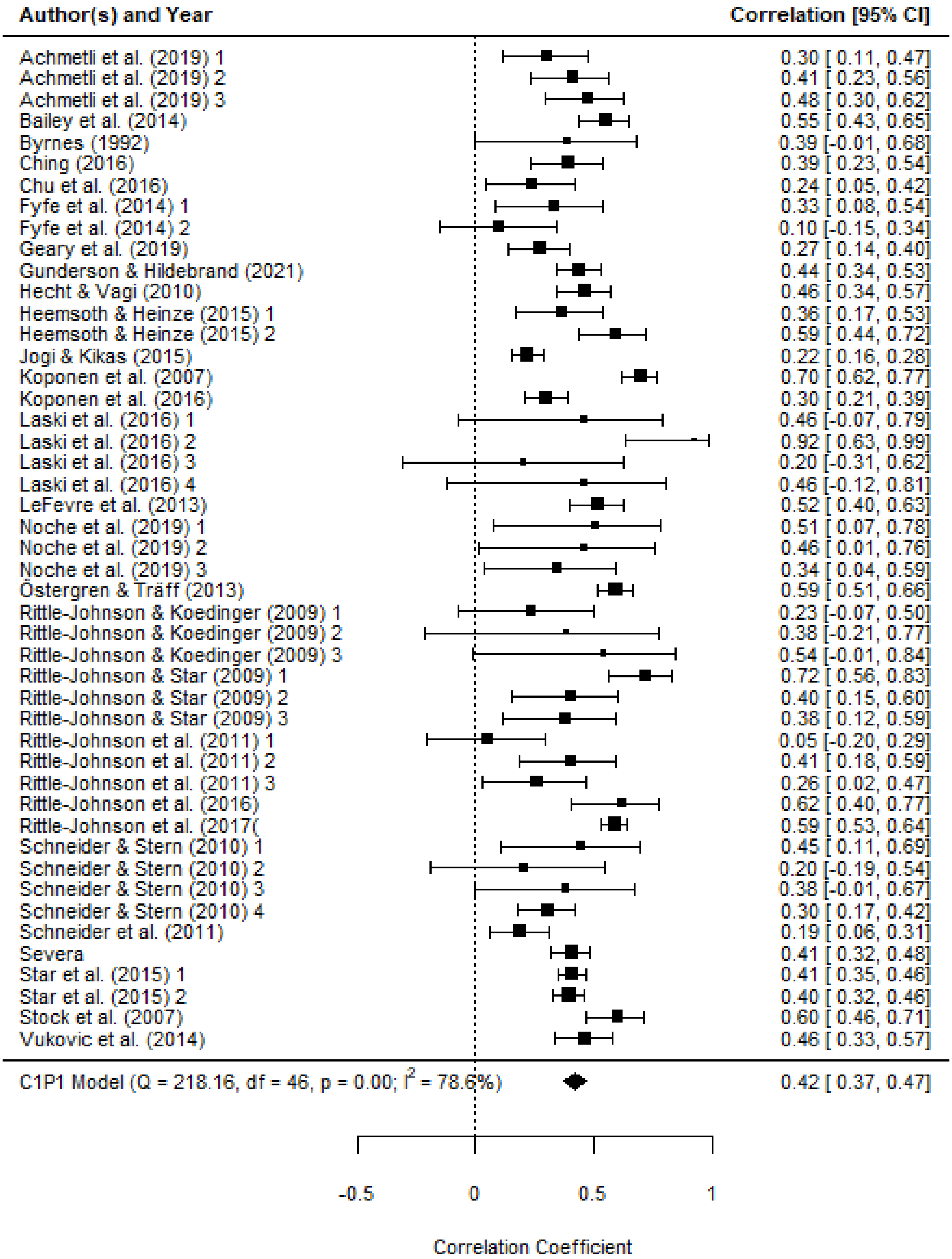


Note. Funnel Plot of the included aggregated bivariate correlations on sample level.

Appendix D: Forest Plots of Included Studies for each Correlation (Study 1)

Figure 41

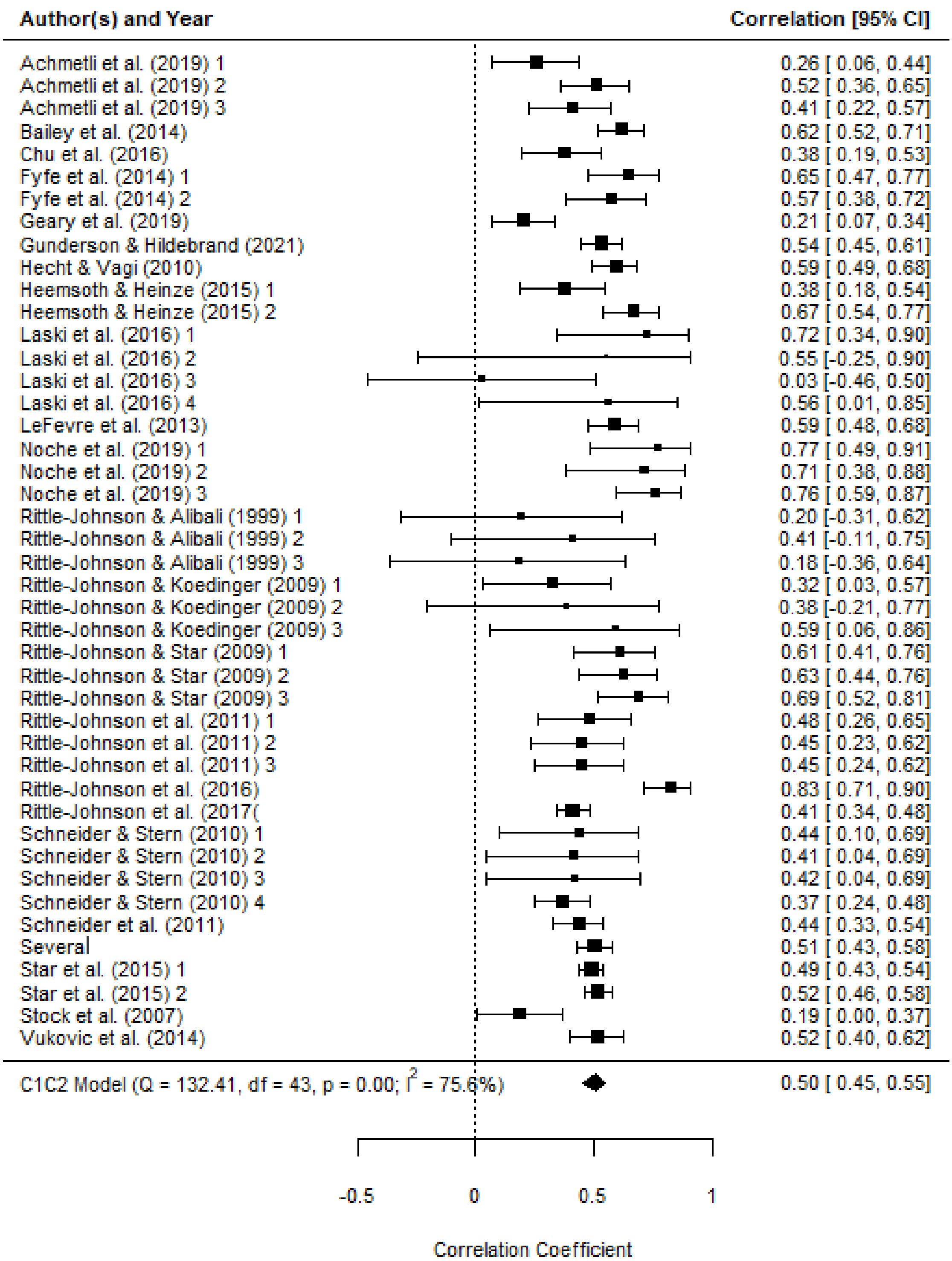
Forest Plot of the Bivariate Correlations Between Conceptual and Procedural Knowledge at T1



Note. Forest Plot of included aggregated bivariate correlations. The digits in the references indicate the subsample based on independent samples in the articles. Several - Aggregated effect size of studies using the same sample (Bailey et al., 2017; Hansen et al., 2015; Hansen et al., 2017; Jordan et al., 2013; Ye et al., 2016).

Figure 42

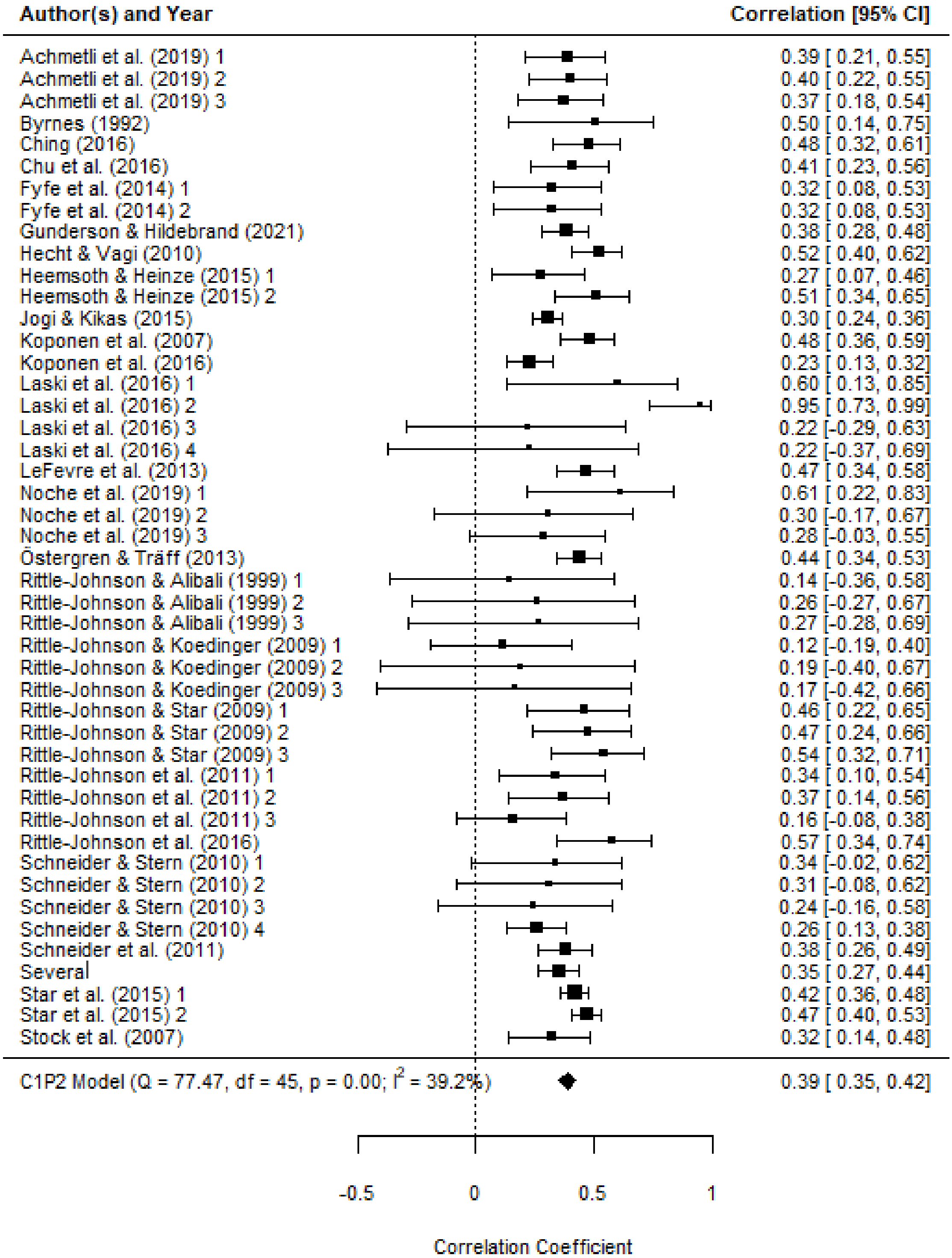
Forest Plot of the Bivariate Correlations Between Conceptual Knowledge at T1 and T2



Note. Forest Plot of included aggregated bivariate correlations. The digits in the references indicate the subsample based on independent samples in the articles. Several - Aggregated effect size of studies using the same sample (Bailey et al., 2017; Hansen et al., 2015; Hansen et al., 2017; Jordan et al., 2013; Ye et al., 2016).

Figure 43

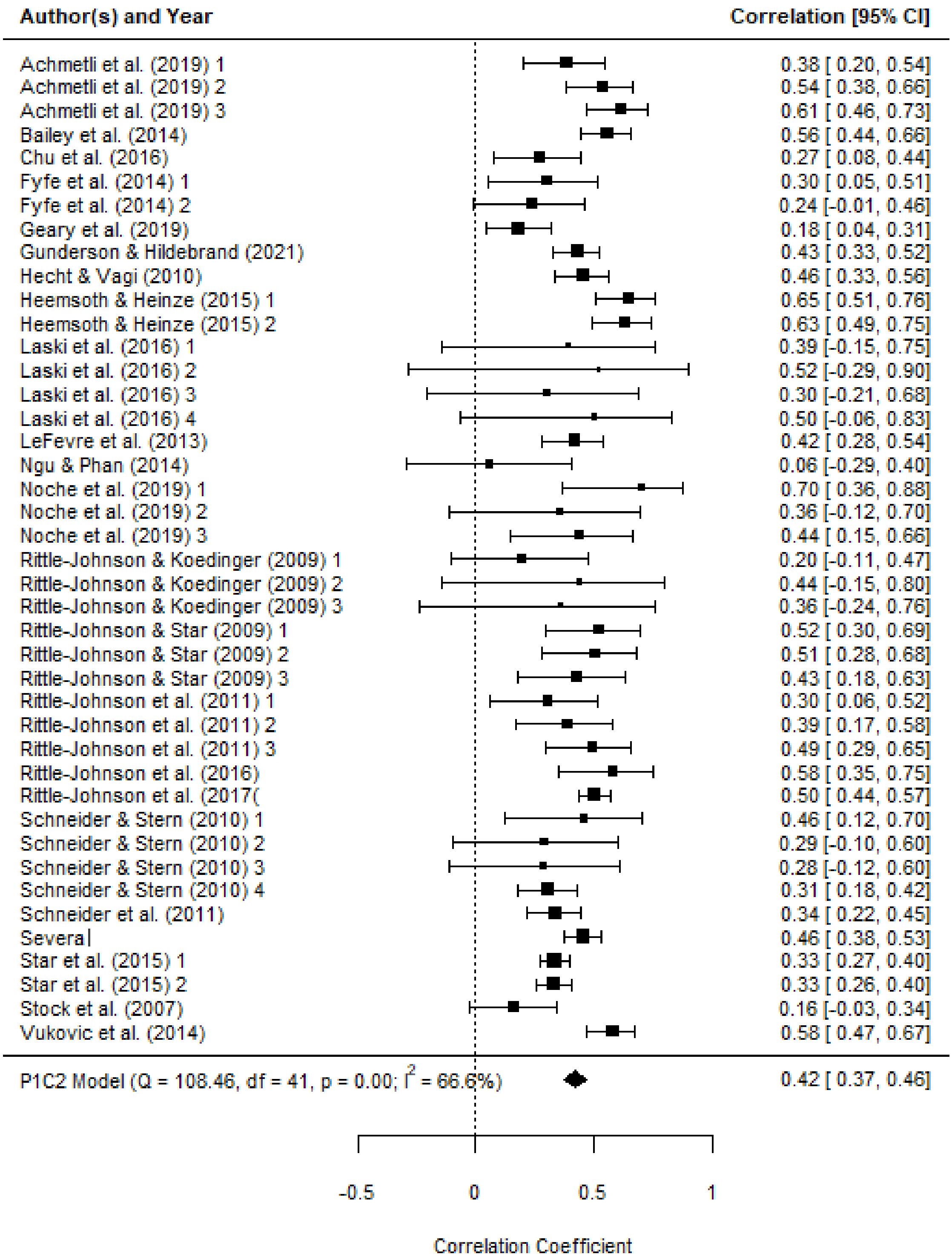
Forest Plot of the Bivariate Correlations Between Conceptual Knowledge at T1 and Procedural Knowledge at T2



Note. Forest Plot of included aggregated bivariate correlations. The digits in the references indicate the subsample based on independent samples in the articles. Several - aggregated effect size of studies using the same sample (Bailey et al., 2017; Hansen et al., 2015; Hansen et al., 2017; Jordan et al., 2013; Ye et al., 2016).

Figure 44

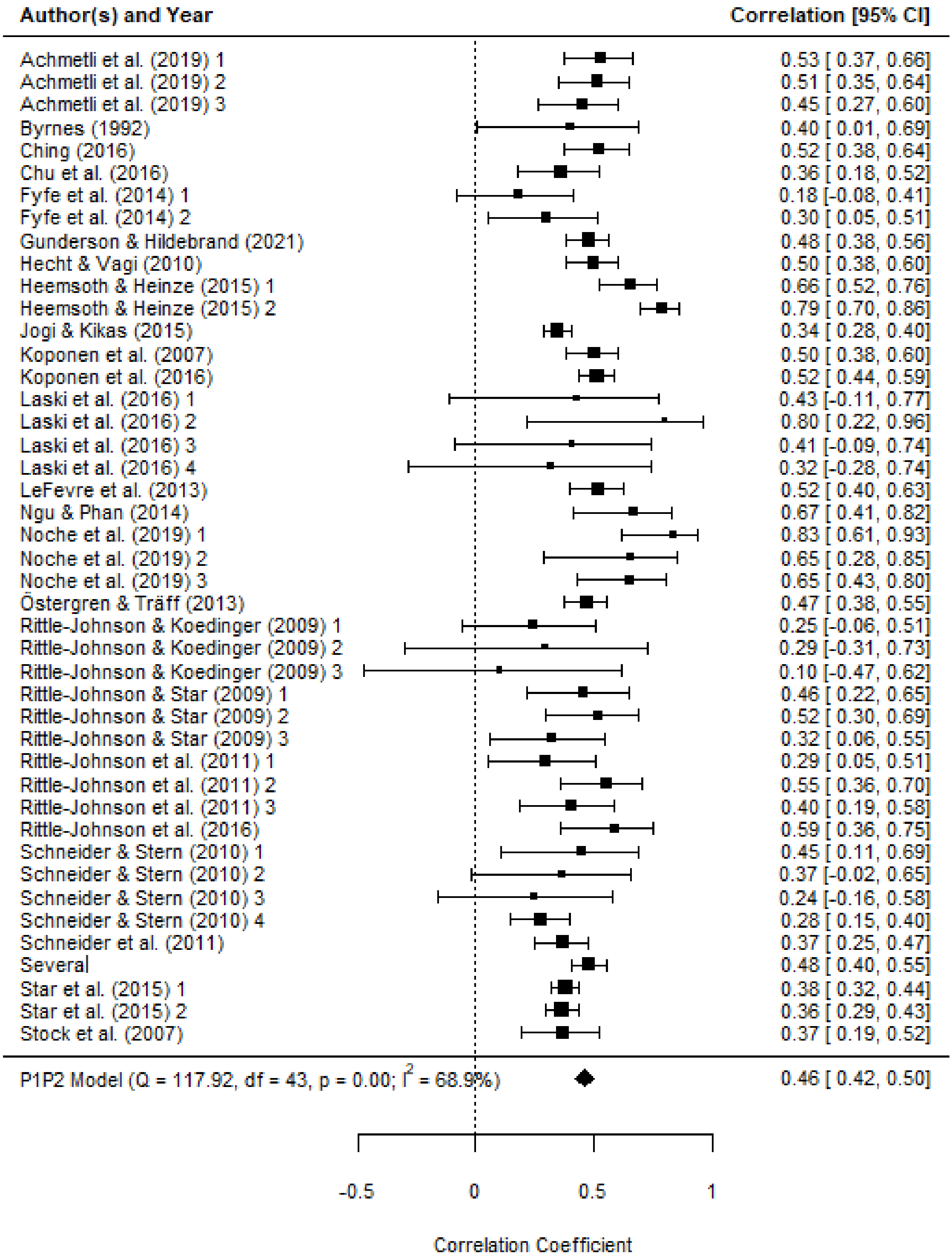
Forest Plot of the Bivariate Correlations Between Procedural Knowledge at T1 and Conceptual Knowledge at T2



Note. Forest Plot of included aggregated bivariate correlations. The digits in the references indicate the subsample based on independent samples in the articles. Several - aggregated effect size of studies using the same sample (Bailey et al., 2017; Hansen et al., 2015; Hansen et al., 2017; Jordan et al., 2013; Ye et al., 2016).

Figure 45

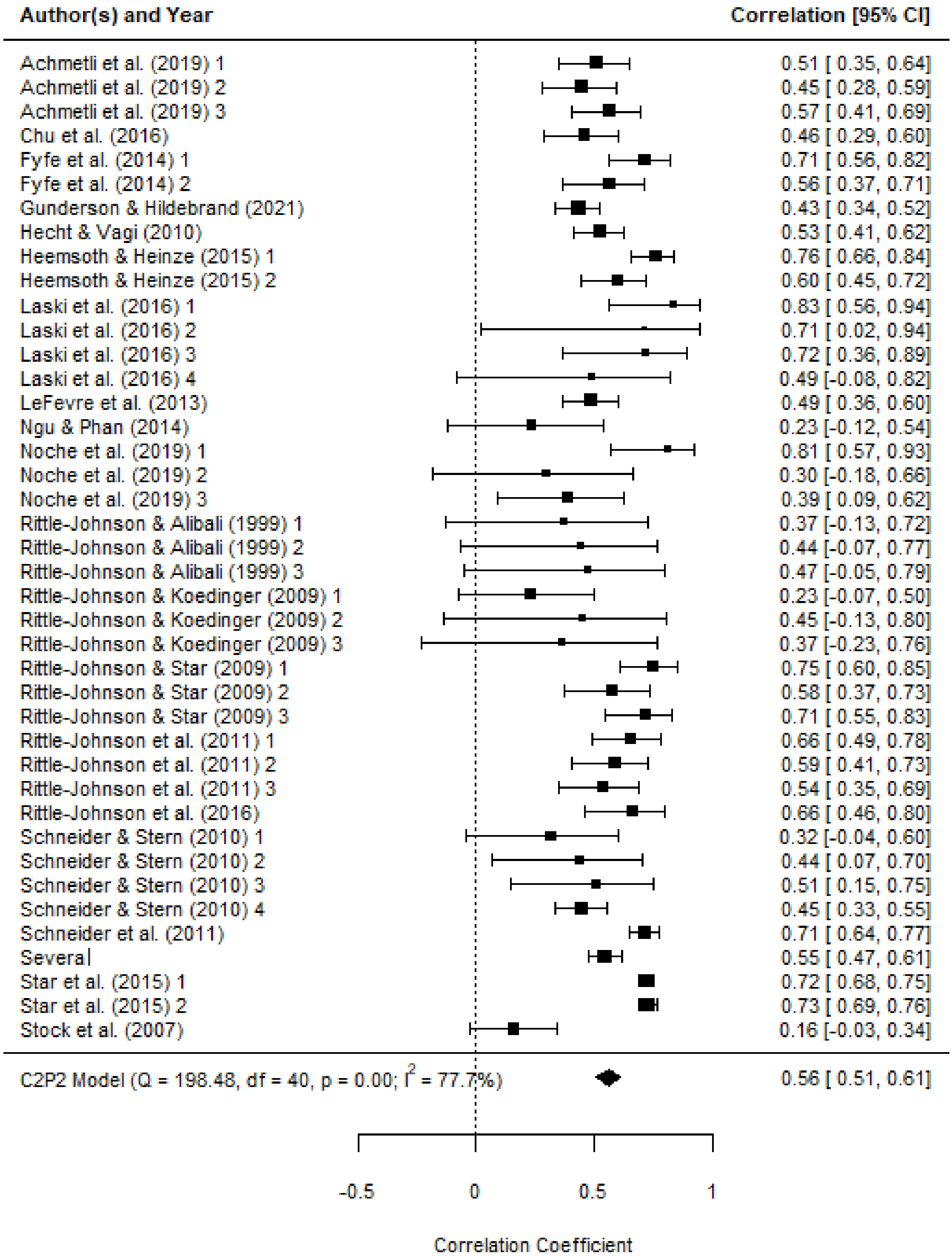
Forest Plot of the Bivariate Correlations Between Procedural Knowledge at T1 and T2



Note. Forest Plot of included aggregated bivariate correlations. The digits in the references indicate the subsample based on independent samples in the articles. Several - aggregated effect size of studies using the same sample (Bailey et al., 2017; Hansen et al., 2015; Hansen et al., 2017; Jordan et al., 2013; Ye et al., 2016).

Figure 46

Forest Plot of the Bivariate Correlations Between Conceptual and Procedural Knowledge at T2



Note. Forest Plot of included aggregated bivariate correlations. The digits in the references indicate the subsample based on independent samples in the articles. Several - aggregated effect size of studies using the same sample (Bailey et al., 2017; Hansen et al., 2015; Hansen et al., 2017; Jordan et al., 2013; Ye et al., 2016).

Appendix E: Risk of Bias Assessment of Included Studies (Study 1)

Table 18

Questions for Risk of Bias Assessment of Included Studies (Study 1)

1	**Baseline Confounding**	**Yes**	**No**	**Not applicable**
1a	Potential baseline confounding was addressed (e.g., by design or adjustments).			
	Rating:			
2	**Sample selection bias**	**Yes**	**No**	**Not applicable**
2a	The drop-out rate is documented and presented.			
2b	The drop-out rate is below 20%.			
2c	Exact drop-out in numbers:			
	Rating:			
3	**Bias in measurement**	**Yes**	**No**	**Not applicable**
3a	The method of measurement was appropriate.			
3b	The reliability of all measurements is mentioned.			
3c	The reported reliability is adequately high.			
	Rating:			
4	**Selective reporting of outcomes**	**Yes**	**No**	**Not applicable**
4a	The numerical result is not likely to have been selected from multiple measurements.			
4b	All six correlations of the four constructs are documented / were delivered.			
	Rating:			
5	**Funding**	**Yes**	**No**	**Not applicable**
5a	The study was not externally and commercially funded.			

Note. Questions of Risk of Bias assessment based on RoB2 (Sterne et al., 2019). Questions are answered for each included study to determine the risk of bias by inclusion in the analysis.

Table 19

Interpretation of Risk of Bias Assessment of Included Studies (Study 1)

Overall Risk of Bias judgement	Interpretation	Criterion
Low risk of bias	The study is comparable to a well-performed randomized trial.	The study is judged to be at **low risk of bias for all domains** for this result.
Moderate risk of bias	The study appears to provide sound evidence for a non-randomized study but cannot be considered comparable to a well-performed randomized trial.	The study is judged to be at **low or moderate risk of bias for all domains.**
Serious risk of bias	The study has one or more critical problems.	The study is judged to be at **serious risk of bias** in at least one domain, but not at critical risk of bias in any domain.
Critical risk of bias	The study is too problematic to provide any helpful evidence and should not be included in any synthesis.	The study is judged to be at **critical risk of bias in at least one domain.**

Note. Interpretation of the Risk of Bias assessment results for each included study following RoB2 (Sterne et al., 2019).

Table 20

Overview of Risk of Bias Assessment of Included Studies (Study 1)

Study	1a	2a	2b	2c [a]	3a	3b	3c	4a	4b	5a
Achmetli et al. (2019)	yes	yes	yes	6,5%	yes	yes	yes	yes	yes	yes
Bailey et al. (2014)	yes	no	n/a	n/a	yes	no	n/a	yes	no	yes
Bailey et al. (2017)	yes	yes	yes	36(517)	yes	yes	yes	yes	yes	yes
Hansen et al. (2015)	yes	yes	yes	75(409)	yes	yes	yes	yes	no	yes
Hansen et al. (2017)	yes	yes	yes	36(517)	yes	yes	yes	yes	no	yes
Jordan et al. (2013)	yes	yes	no	124(481)	yes	no	n/a	yes	no	yes
Ye et al. (2016)	yes	yes	yes	36(517)	yes	yes	yes	yes	yes	yes
Byrnes (1992)	no	no	n/a	n/a	yes	no	n/a	yes	no	no
Ching (2016)	yes	yes	n/a	n/a	yes	yes	yes	yes	no	yes
Chu et al. (2016)	yes	yes	no	53(153)	yes	no	n/a	yes	yes	yes
Cowan et al. (2011)	yes	yes	yes	10(269)	yes	no	n/a	yes	yes	yes
Fyfe et al. (2014)	yes	yes	yes	14(183)	yes	no	yes	yes	yes	yes
Geary et al. (2019)	yes	yes	yes	35(232)	yes	no	n/a	yes	no	yes
Gunderson & Hildebrand (2021)	yes	yes	yes	15(612)	yes	yes	yes	yes	yes	yes
Hecht & Vagi (2010)	yes	yes	no	79(260)	yes	no	?	yes	yes	yes
Heemsoth & Heinze (2015)	yes	yes	yes	32(206)	yes	yes	no	yes	yes	yes
Jogi & Kikas (2015)	yes	no	n/a	n/a	yes	yes	yes	yes	no	yes
Koponen et al. (2007)	yes	yes	yes	29(207)	yes	yes	yes	yes	no	yes

Table 20 (continued)

Study	1a	2a	2b	2c [a]	3a	3b	3c	4a	4b	5a
Koponen et al. (2016)	yes	n/a	n/a	n/a	yes	yes	yes	yes	no	yes
Laski et al (2016)	yes	yes	no	97(150)	yes	no	n/a	yes	yes	yes
LeFevre et al. (2013)	yes	yes	yes	3(160)	yes	yes	yes	yes	yes	yes
Ngu & Phan (2016)	yes	yes	yes	1(63)	yes	no	n/a	yes	no	yes
Noche et al. (2019)	yes	yes	no	103(267)	yes	yes	yes	yes	yes	yes
Östergren & Träff (2013)	no	yes	yes	10(315)	yes	yes	yes	yes	no	yes
Rittle-Johnson & Koedinger (2009)	yes	yes	yes	11(88)/ 3(26)	yes	no	n/a	yes	yes	no
Rittle-Johnson & Star (2009)	yes	yes	yes	8(162)	yes	no	n/a	yes	yes	yes
Rittle-Johnson et al. (2011)	yes	yes	no	52(250)	yes	no	n/a	yes	yes	yes
Rittle-Johnson et al. (2016)	yes	yes	yes	21(201)	yes	no	n/a	yes	yes	yes
Rittle-Johnson et al. (2017)	yes	yes	no	254	yes	yes	yes	yes	no	yes
Schneider & Stern (2010)	yes	yes	yes	9/26	yes	no	yes	yes	yes	yes
Schneider et al. (2011)	yes	yes	yes	11/ 21	yes	no	n/a	yes	yes	yes
Star et al. (2015)	yes	yes	no	65(517)	yes	yes	yes	yes	yes	yes
Stock et al. (2007)	yes	yes	n/a	n/a	yes	yes	yes	yes	yes	yes
Vukovic et al. (2014)	yes	n/a	n/a	n/a	yes	yes	yes	yes	no	yes
Wasiu & Abiola (2019)	yes	no	n/a	n/a	yes	no	n/a	yes	yes	yes

Note. Overview of results in the Risk of Bias assessment for each study. The columns refer to the questions about the Risk of Bias assessment (see Table 4). yes - The respective question is answered with a yes for this study; no - The respective question is answered with a no for this study; n/a - The question does not apply to this study.

[a] – The exact drop-out rate in numbers with the total number of participants in brackets.

Table 21

Result of Risk of Bias Assessment for Included Studies (Study 1)

Study	Baseline Confounding	Sample selection bias	Bias in measurement	Selective reporting of outcomes	External commercial funding	Overall rating
Achmetli et al. (2019)	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Bailey et al. (2014)	Low risk	Moderate risk	Moderate risk	Moderate risk	Low risk	Moderate risk
Bailey et al. (2017)	Low risk	Low risk	Low risk	Low risk	Low risk	Moderate risk
Hansen et al. (2015)	Low risk	Low risk	Low risk	Moderate risk	Low risk	Moderate risk
Hansen et al. (2017)	Low risk	Low risk	Low risk	Moderate risk	Low risk	Moderate risk
Jordan et al. (2013)	Low risk	Moderate risk	Moderate risk	Moderate risk	Low risk	Moderate risk
Ye et al. (2016)	Low risk	Low risk	Low risk	Low risk	Low risk	Moderate risk
Byrnes (1992)	Moderate risk	Moderate risk	Moderate risk	Moderate risk	Low risk	Moderate risk
Ching (2016)	Low risk	Low risk	Low risk	Moderate risk	Low risk	Moderate risk
Chu et al. (2016)	Low risk	Moderate risk	Moderate risk	Low risk	Low risk	Moderate risk
Cowan et al. (2011)	Low risk	Low risk	Moderate risk	Low risk	Low risk	Moderate risk
Fyfe et al. (2014)	Low risk	Low risk	Moderate risk	Low risk	Low risk	Moderate risk
Geary et al. (2019)	Low risk	Low risk	Moderate risk	Moderate risk	Low risk	Moderate risk
Gunderson & Hildebrand (2021)	Low risk	Low risk	Low risk	Low risk	Low risk	Moderate risk
Hecht & Vagi (2010)	Low risk	Moderate risk	Moderate risk	Low risk	Low risk	Moderate risk
Heemsoth & Heinze (2015)	Low risk	Low risk	Moderate risk	Low risk	Low risk	Moderate risk
Jogi & Kikas (2015)	Low risk	Moderate risk	Low risk	Moderate risk	Low risk	Moderate risk
Koponen et al. (2007)	Low risk	Low risk	Low risk	Moderate risk	Low risk	Moderate risk

Table 21 (continued)

Study	Baseline Confounding	Sample selection bias	Bias in measurement	Selective reporting of outcomes	External commercial funding	Overall rating
Koponen et al. (2016)	Low risk	n/a	Low risk	Moderate risk	Low risk	Moderate risk
Laski et al. (2016)	Low risk	Moderate risk	Moderate risk	Low risk	Low risk	Moderate risk
LeFevre et al. (2013)	Low risk	Low risk	Low risk	Low risk	Low risk	Moderate risk
Ngu & Phan (2016)	Low risk	Low risk	Moderate risk	Moderate risk	Low risk	Moderate risk
Noche et al. (2019)	Low risk	Moderate risk	Low risk	Low risk	Low risk	Moderate risk
Östergren & Träff (2013)	Moderate risk	Low risk	Low risk	Low risk	Low risk	Moderate risk
Rittle-Johnson & Koedinger (2009)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	Moderate risk
Rittle-Johnson & Star (2009)	Low risk	Low risk	Moderate risk	Low risk	Low risk	Moderate risk
Rittle-Johnson et al. (2011)	Low risk	Moderate risk	Moderate risk	Low risk	Low risk	Moderate risk
Rittle-Johnson et al. (2016)	Low risk	Low risk	Moderate risk	Low risk	Low risk	Moderate risk
Rittle-Johnson et al. (2017)	Low risk	Moderate risk	Low risk	Moderate risk	Low risk	Moderate risk
Schneider & Stern (2010)	Low risk	Low risk	Moderate risk	Low risk	Low risk	Moderate risk
Schneider et al. (2011)	Low risk	Low risk	Moderate risk	Low risk	Low risk	Moderate risk
Star et al. (2015)	Low risk	Moderate risk	Low risk	Low risk	Low risk	Moderate risk
Stock et al. (2007)	Low risk	Low risk	Low risk	Low risk	Low risk	Moderate risk
Vukovic et al. (2014)	Low risk	n/a	Low risk	Moderate risk	Low risk	Moderate risk
Wasiu & Abiola (2019)	Low risk	Moderate risk	Moderate risk	Low risk	Low risk	Moderate risk

Note. Overview of results in the Risk of Bias assessment for each study. The columns refer to the interpretation of the Risk of Bias assessment (see Table 19).

Appendix F: Included Studies (Study 2)

Aladé, F., & Nathanson, A. I. (2016). What preschoolers bring to the show: The relation between viewer characteristics and children's learning from educational television. *Media Psychology*, *19*(3), 406–430. https://doi.org/10.1080/15213269.2015.1054945

Alao, S., & Guthrie, J. T. (1999). Predicting conceptual understanding with cognitive and motivational variables. *The Journal of Educational Research*, *92*(4), 243–254. https://doi.org/10.1080/00220679909597602

Alexander, P. A., Kulikowich, J. M., & Schulze, S. K. (1994). The influence of topic knowledge, domain knowledge, and interest on the comprehension of scientific exposition. *Learning and Individual Differences*, *6*(4), 379–397. https://doi.org/10.1016/1041-6080(94)90001-9

Alexander, P. A., Jetton, T. L., & Kulikowich, J. M. (1995). Interrelationship of knowledge, interest, and recall: Assessing a model of domain learning. *Journal of Educational Psychology*, *87*(4), 559–575.

Almeida, J., O'Brien, K. H., Gironda, C. M., & Gross, E. B. (2017). Development, implementation, and evaluation of a comprehensive course on suicide in a master's of social work program. *Journal of Social Work Education*, *53*(4), 727–736.

Anderman, E. M., & Johnston, J. (1994). Achievement Goals and Current Events Knowledge. In *Annual Meeting of the American Educational Research Association*.

Anmarkrud, Ø., & Bråten, I. (2009). Motivation for reading comprehension. *Learning and Individual Differences*, *19*(2), 252–256. https://doi.org/10.1016/j.lindif.2008.09.002

Arens, A. K., Marsh, H. W., Craven, R. G., Yeung, A. S., Randhawa, E., & Hasselhorn, M. (2016). Math self-concept in preschool children: Structure, achievement relations, and generalizability across gender. *Early Childhood Research Quarterly*, *36*, 391–403. https://doi.org/10.1016/j.ecresq.2015.12.024

Bailey, J. M., Lombardi, D., Cordova, J. R., & Sinatra, G. M. (2017). Meeting students halfway: Increasing self-efficacy and promoting knowledge change in astronomy. *Physical Review Physics Education Research*, *13*(2), Article 020140. https://doi.org/10.1103/PhysRevPhysEducRes.13.020140

Barber, M. E. (2010). *Students' prior knowledge, ability, motivation, test anxiety, and course engagement as predictors of learning in community college psychology courses* (Publication No. 3496127) [Doctoral Dissertation, University of Florida]. ProQuest LLC.

Bleicher, R. E. (2007). Nurturing confidence in preservice elementary science teachers. *Journal of Science Teacher Education*, *18*(6), 841–860. https://doi.org/10.1007/s10972-007-9067-2

Bråten, I., Anmarkrud, Ø., Brandmo, C., & Strømsø, 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. https://doi.org/10.1016/j.learninstruc.2013.11.002

Brunstein, J. C., & Glaser, C. (2011). Testing a path-analytic mediation model of how self-regulated writing strategies improve fourth graders' composition skills: A randomized controlled trial. *Journal of Educational Psychology*, *103*(4), 922–938. https://doi.org/10.1037/a0024622

Chae. S. E. (2011). *Contributions of prior knowledge, motivation and strategies to Korean college students' L2 writing development* (http://hdl.handle.net/1903/12261) [Doctoral Dissertation, University of Maryland]. Digital Repository at the University of Maryland.

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. https://doi.org/10.1016/j.compedu.2011.09.012

Conradty, C., & Bogner, F. X. (2016). Hypertext or textbook: effects on motivation and gain in knowledge. *Education Sciences*, *6*(3), Article 29. https://doi.org/10.3390/educsci6030029

Cross, D., Lester, L., Pearce, N., Barnes, A., & Beatty, S. (2018). A group randomized controlled trial evaluating parent involvement in whole-school actions to reduce bullying. *The Journal of Educational Research*, *111*(3), 255–267. https://doi.org/10.1080/00220671.2016.1246409

Dinsmore, D. L., Alexander, P. A., & Loughlin, S. M. (2008). The impact of new learning environments in an engineering design course. *Instructional Science*, *36*, 375–393. https://doi.org/10.1007/s11251-008-9061-x

Dinsmore, D. L., & Zoellner, B. P. (2018). The relation between cognitive and metacognitive strategic processing during a science simulation. *British Journal of Educational Psychology*, *88*(1), 95–117. https://doi.org/10.1111/bjep.12177

Guerin, R. J., Toland, M. D., Okun, A. H., Rojas-Guyler, L., & Bernard, A. L. (2018). Using a modified theory of planned behavior to examine adolescents' workplace safety and health knowledge, perceptions, and behavioral intention: A structural equation modeling approach. *Journal of Youth and Adolescence*, *47*, 1595–1610. https://doi.org/10.1007/s10964-018-0847-0

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. https://doi.org/10.1080/01443410701413753

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

Hambrick, D. Z., Pink, J. E., Meinz, E. J., Pettibone, J. C., & Oswald, F. L. (2008). The roles of ability, personality, and interests in acquiring current events knowledge: A longitudinal study. *Intelligence*, *36*(3), 261–278.

https://doi.org/10.1016/j.intell.2007.06.004

Haverback, H. R. (2007). *Pre-service reading teachers as tutors: An examination of efficacy and content knowledge* (Publication No. 3260381) [Doctoral Dissertation, University of Maryland]. ProQuest LLC.

Holmes, V. (2012). New Digital Energy Game, the Use of Games to Influence Attitudes, Interests, and Student Achievement in Science. *Online Submission*.

Hwang, G. J., Sung, H. Y., & Chang, H. (2017). Effects of concept-mapping-based interactive e-books on active and reflective-style students' learning performances in junior high school law courses. *Interactive Learning Environments*, *25*(7), 877–888. https://doi.org/10.1080/10494820.2016.1224253

Höft, L., & Bernholt, S. (2019). Longitudinal couplings between interest and conceptual understanding in secondary school chemistry: an activity-based perspective. *International Journal of Science Education*, *41*(5), 607–627. https://doi.org/10.1080/09500693.2019.1571650

Joiner, R., Iacovides, J., Owen, M., Gavin, C., Clibbery, S., Darling, J., & Drew, B. (2011). Digital games, gender and learning in engineering: Do females benefit as much as males?. *Journal of Science Education and Technology*, *20*, 178–185. https://doi.org/10.1007/s10956-010-9244-5

Kane, M. J., Smeekens, B. A., Von Bastian, C. C., Lurquin, J. H., Carruth, N. P., & Miyake, A. (2017). A combined experimental and individual-differences investigation into mind wandering during a video lecture. *Journal of Experimental Psychology: General*, *146*(11), 1649–1674. https://doi.org/10.1037/xge0000362

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

Liebfreund, M. D., & Conradi, K. (2016). Component skills affecting elementary students' informational text comprehension. *Reading and Writing*, *29*, 1141–1160.

https://doi.org/10.1007/s11145-016-9629-9

Linnenbrink-Garcia, L., Pugh, K. J., Koskey, K. L., & 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. https://doi.org/10.1080/00220973.2011.559491

List, A., Stephens, L. A., & Alexander, P. A. (2019). Examining interest throughout multiple text use. *Reading and Writing*, *32*, 307–333. https://doi.org/10.1007/s11145-018-9863-4

Liu, Y., & Hou, S. (2018). Potential reciprocal relationship between motivation and achievement: A longitudinal study. *School Psychology International*, *39*(1), 38–55.

Marsh, H. W., Byrne, B. M., & Yeung, A. S. (1999). Causal ordering of academic self-concept and achievement: Reanalysis of a pioneering study and... *Educational Psychologist*, *34*(3), 155–167. https://doi.org/10.1207/s15326985ep3403_2

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, *76*(2), 397–416. https://doi.org/10.1111/j.1467-8624.2005.00853.x

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*, 197–216. https://doi.org/10.1007/s11409-009-9045-5

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

Navratil, S. D., Kühl, T., & Heidig, S. (2018). Why the cells look like that-the influence of learning with emotional design and elaborative interrogations. *Frontiers in*

Psychology, *9*, Article 1653. https://doi.org/10.3389/fpsyg.2018.01653

Nussbaum, E. M., Owens, M. C., Sinatra, G. M., Rehmat, A. P., Cordova, J. R., Ahmad, S. Harris, F. C., Dascalu, S. M. (2015). Losing the Lake: Simulations to promote gains in student knowledge and interest about climate change. *International Journal of Environmental and Science Education*, *10*(6), 789–811.

Pekrun, R., Hall, N. C., Goetz, T., & Perry, R. P. (2014). Boredom and academic achievement: Testing a model of reciprocal causation. *Journal of Educational Psychology*, *106*(3), 696–710. https://doi.org/10.1037/a0036006

Pinger, P., Rakoczy, K., Besser, M., & Klieme, E. (2018). Implementation of formative assessment-effects of quality of programme delivery on students' mathematics achievement and interest. *Assessment in Education: Principles*, *Policy & Practice*, *25*(2), 160–182. https://doi.org/10.1080/0969594X.2016.1170665

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 and Practice*, *14*(5), 1997–2012.

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.

Schneider, S., Häßler, A., Habermeyer, T., Beege, M., & Rey, G. D. (2019). The more human, the higher the performance? Examining the effects of anthropomorphism on learning with media. *Journal of Educational Psychology*, *111*(1), 57–72. https://doi.org/10.1037/edu0000273

Seaton, M., Parker, P., Marsh, H. W., Craven, R. G., & Yeung, A. S. (2014). The reciprocal relations between self-concept, motivation and achievement: Juxtaposing academic self-concept and achievement goal orientations for mathematics success. *Educational Psychology*, *34*(1), 49–72. https://doi.org/10.1080/01443410.2013.825232

Seifried, E., Eckert, C., & Spinath, B. (2018). Optional Learning Opportunities: Who Seizes Them and What Are the Learning Outcomes? *Teaching of Psychology*, *45*(3), 246–250.

Shanley, C. J. (2014). *1+ 1 is not always 2: Variation in the relations between mathematics self-efficacy development and longitudinal mathematics achievement growth* (Publication No. 3673121) [Doctoral Dissertation, University of Oregon]. ProQuest LLC.

Skuballa, I. T., Xu, K. M., & Jarodzka, H. (2019). The impact of co-actors on cognitive load: when the mere presence of others makes learning more difficult. *Computers in Human Behavior*, *101*, 30–41. https://doi.org/10.1016/j.chb.2019.06.016

Son, E. (2015). *Math achievement and self-efficacy of linguistically and ethnically diverse high school students: Their relationships with english reading and native language proficiency* (Publication No. 10014546) [Doctoral dissertation, University of Southern California]. ProQuest LLC.

Strømsø, H. I., Bråten, I., & Britt, M. A. (2010). Reading multiple texts about climate change: The relationship between memory for sources and text comprehension. *Learning and Instruction*, *20*(3), 192–204. https://doi.org/10.1016/j.learninstruc.2009.02.001

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

Torres, W. J., & Beier, M. E. (2018). Adult development in the wild: The determinants of autonomous learning in a Massive Open Online Course. *Learning and Individual Differences*, *65*, 207–217. https://doi.org/10.1016/j.lindif.2018.06.003

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

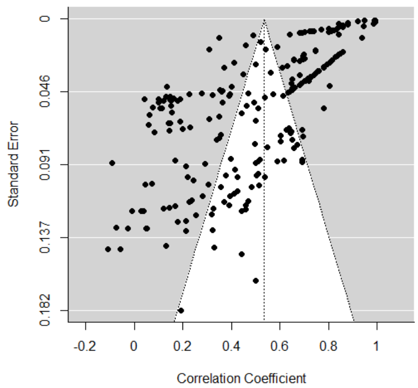
Watts, T. W., Duncan, G. J., Chen, M., Claessens, A., Davis-Kean, P. E., Duckworth, K., Engel, M., Siegler, R., & Susperreguy, M. I. (2015). The role of mediators in the development of longitudinal mathematics achievement associations. *Child Development*, *86*(6), 1892–1907. https://doi.org/10.1111/cdev.12416

Appendix G: Funnel Plots for the Overall Correlation Between Prior Knowledge and Learning Outcomes (Study 2)

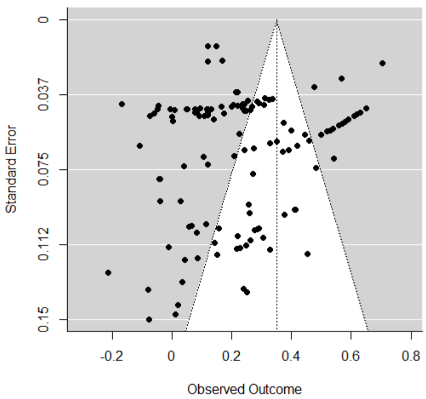
Figure 47

Funnel Plots for Overall Correlation Between Prior Knowledge and Learning Outcomes

Correlations between prior knowledge and learning outcomes



Funnel plots for interest



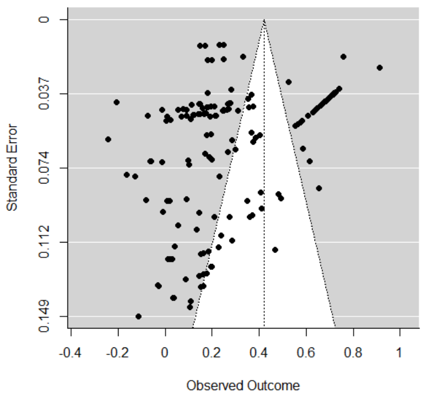
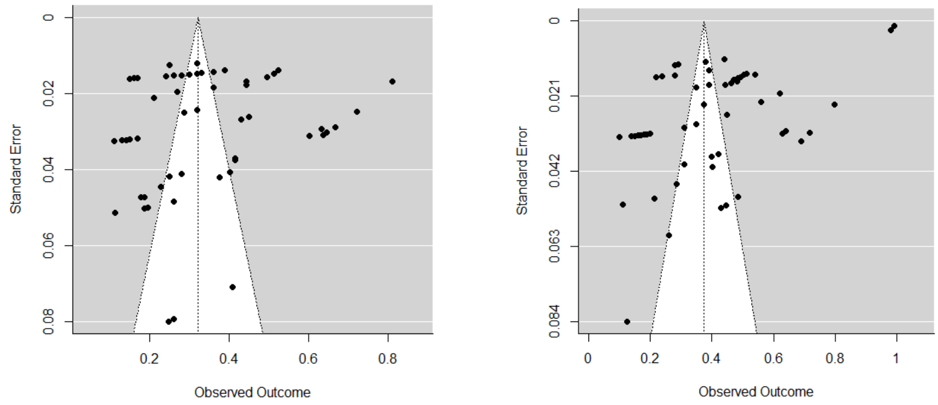
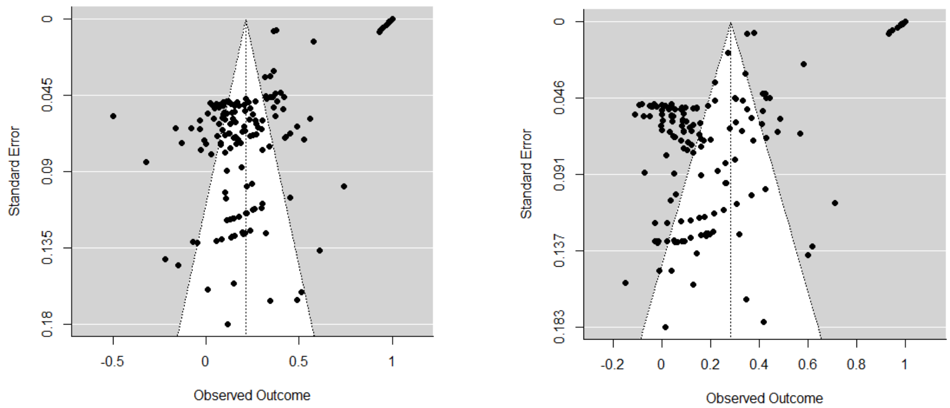


Figure 47 (continued)

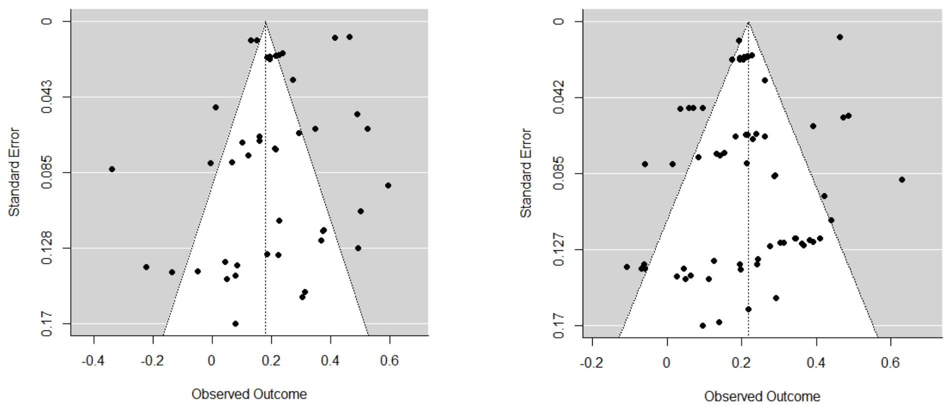
Funnel plots for self-concept



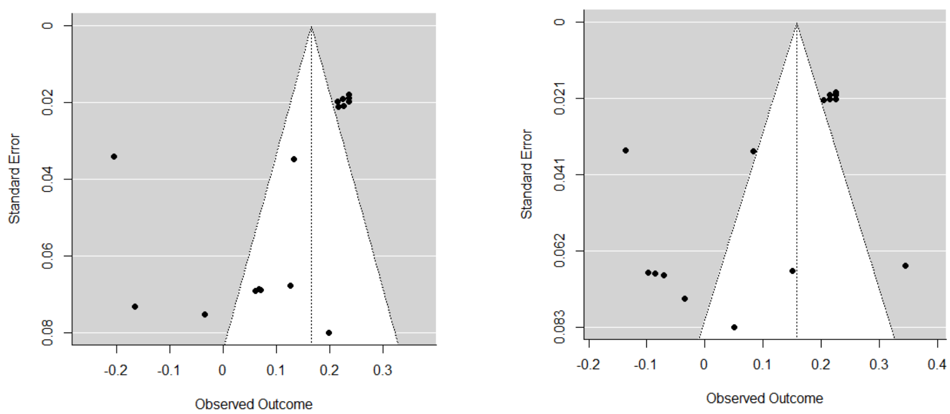
Funnel plots for self-efficacy



Funnel plots for intrinsic motivation



Funnel plots for extrinsic motivation



Note. b_1-paths (left), b_2-paths (right), Separated by Mediators.

Appendix H: Included Studies (Study 3)

Agostinho, S., Tindall-Ford, S., Ginns, P., Howard, S. J., Leahy, W., & Paas, F. (2015). Giving learning a helping hand: Finger tracing of temperature graphs on an iPad. *Educational Psychology Review*, *27*, 427–443. https://doi.org/10.1007/s10648-015-9315-5

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps. *Journal of Educational Psychology*, *95*(4), 774–783. https://doi.org/10.1037/0022-0663.95.4.774

Bainbridge, K., Shute, V., Rahimi, S., Liu, Z., Slater, S., Baker, R. S., & D'Mello, S. K. (2022). Does embedding learning supports enhance transfer during game-based learning?. *Learning and Instruction*, *77*, Article 101547. https://doi.org/10.1016/j.learninstruc.2021.101547

Barenberg, J., Berse, T., Reimann, L., & Dutke, S. (2021). Testing and transfer: Retrieval practice effects across test formats in English vocabulary learning in school. *Applied Cognitive Psychology*, *35*(3), 700–710. https://doi.org/10.1002/acp.3796

Barron, B. (2000). Problem solving in video-based microworlds: Collaborative and individual outcomes of high-achieving sixth-grade students. *Journal of Educational Psychology*, *92*(2), 391–398. https://doi.org/10.1037/0022-0663.92.2.391

Basile, C. G. (2000). Environmental education as a catalyst for transfer of learning in young children. *The Journal of Environmental Education*, *32*(1), 21–27. https://doi.org/10.1080/00958960009598668

Beeson, G. W. (1981). Influence of knowledge context on the learning of intellectual skills. *American Educational Research Journal*, *18*(3), 363–379.

Beilstein, S. O. (2019). *Supporting children's conceptual understanding of fractions with manipulatives and gesture*. (http://hdl.handle.net/2142/104770) [Doctoral dissertation, University of Illinois]. Shereen Oca Beilstein.

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. https://doi.org/10.1111/bjep.12062

Bergey, B. W., Cromley, J. G., & Newcombe, N. S. (2015). Teaching high school biology students to coordinate text and diagrams: Relations with transfer, effort, and spatial skill. *International Journal of Science Education*, *37*(15), 2476–2502. https://doi.org/10.1080/09500693.2015.1082672

Bottge, B. A. (1999). Effects of contextualized math instruction on problem solving of average and below-average achieving students. *The Journal of Special Education*, *33*(2), 81–92. https://doi.org/10.1177/002246699903300202

Boucheix, J. M., & Guignard, H. (2005). What animated illustrations conditions can improve technical document comprehension in young students? Format, signaling and control of the presentation. *European Journal of Psychology of Education*, *20*, 369–388. https://doi.org/10.1007/BF03173563

Britt, M. A., & Aglinskas, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction*, *20*(4), 485–522. https://doi.org/10.1207/S1532690XCI2004_2

Carbonneau, K. J., & Marley, S. C. (2015). Instructional guidance and realism of manipulatives influence preschool children's mathematics learning. *The Journal of Experimental Education*, *83*(4), 495–513. https://doi.org/10.1080/00220973.2014.989306

Carnine, D., Stevens, C., Clements, J., & Kame'enui, E. J. (1982). Effects of facultative questions and practice on intermediate students' understanding of character motives. *Journal of Reading Behavior*, *14*(2), 179–190.

Casey, B., Erkut, S., Ceder, I., & Young, J. M. (2008). Use of a storytelling context to improve girls' and boys' geometry skills in kindergarten. *Journal of Applied Developmental*

Psychology, *29*(1), 29–48. https://doi.org/10.1016/j.appdev.2007.10.005

Chen, C. L. (2018). *The effects of second-language repeated reading on reading comprehension and incidental vocabulary acquisition*. (Publication No. 10974542). [Doctoral dissertation, Teachers College, Columbia University.] ProQuest LLC.

Chen, Z. (1999). Schema induction in children's analogical problem solving. *Journal of Educational Psychology*, *91*(4), 703–715. https://doi.org/10.1037/0022-0663.91.4.703

Cheng, R. (2019). *Effects of teaching text structure in science text reading: A study among chinese middle school students* (Publication No. 13865151). [Doctoral dissertation, Columbia University]. ProQuest LLC.

Ching, B. H. H., & Wu, X. (2019). Concreteness fading fosters children's understanding of the inversion concept in addition and subtraction. *Learning and Instruction*, *61*, 148–159. https://doi.org/10.1016/j.learninstruc.2018.10.006

Chiu, T. K., & Mok, I. A. (2017). Learner expertise and mathematics different order thinking skills in multimedia learning. *Computers & Education*, *107*, 147–164. https://doi.org/10.1016/j.compedu.2017.01.008

Clarke, A. M., Cooper, G. M., & Loudon, E. H. (1969). A set to establish equivalence relations in pre-school children. *Journal of Experimental Child Psychology*, *8*(1), 180–189. https://doi.org/10.1016/0022-0965(69)90039-3

Clements, D. H., Sarama, J., Baroody, A. J., Joswick, C., & Wolfe, C. B. (2019). Evaluating the efficacy of a learning trajectory for early shape composition. *American Educational Research Journal*, *56*(6), 2509–2530. https://doi.org/10.3102/0002831219842788

Clifton, M. A. (1985). Practice variability and children's motor behavior. *Perceptual and Motor Skills*, *60*(2), 471–476. https://doi.org/10.2466/pms.1985.60.2.471

Colón-Acosta, N. (2019). *Learning to code: Effects of programming modality in a game-based learning environment* (Publication No. 13865235). [Doctoral dissertation, Columbia University]. ProQuest LLC.

Cramer, K. A., & Post, T. R. (2002). Initial fraction learning by fourth-and fifth-grade students: A comparison of the effects of using commercial curricula with the effects of using the rational number project curriculum. *Journal for Research in Mathematics Education*, *33*(2), 111–144. https://doi.org/10.2307/749646

Cromley, J. G., Bergey, B. W., Fitzhugh, S., Newcombe, N., Wills, T. W., Shipley, T. F., & Tanaka, J. C. (2013). Effects of three diagram instruction methods on transfer of diagram comprehension skills: The critical role of inference while learning. *Learning and Instruction*, *26*, 45–58. https://doi.org/10.1016/j.learninstruc.2013.01.003

Cromley, J. G., Perez, T. C., Fitzhugh, S. L., Newcombe, N. S., Wills, T. W., & Tanaka, J. C. (2013). Improving students' diagram comprehension with classroom instruction. *The Journal of Experimental Education*, *81*(4), 511–537. https://doi.org/10.1080/00220973.2012.745465

Day, J. D., & Cordón, L. A. (1993). Static and dynamic measures of ability: An experimental comparison. *Journal of Educational Psychology*, *85*(1), 75–82. https://doi.org/10.1037/0022-0663.85.1.75

DeMarinis, M. D. (2011). *An interactive geometry program and its effect on elementary students' achievement and understanding of geometry: A comparative study* (Publication No. 3484225). [Doctoral Dissertation, Teachers College, Columbia University]. ProQuest LLC.

Douvis, S. J. (2005). Variable practice in learning the forehand drive in tennis. *Perceptual and Motor Skills*, *101*(2), 531–545. https://doi.org/10.2466/pms.101.2.531-545

Du, X., & Zhang, Q. (2019). Tracing worked examples: Effects on learning in geometry. *Educational Psychology*, *39*(2), 169–187. https://doi.org/10.1080/01443410.2018.1536256

Durnin, J. H., Perrone, A. E., & MacKay, L. (1997). Teaching problem solving processes in elementary school mathematics. *Journal of Structural Learning & Intelligent Systems*, *13*(1), 53–69.

Fang, S. C., Hsu, Y. S., & Hsu, W. H. (2016). Effects of explicit and implicit prompts on students' inquiry practices in computer-supported learning environments in high school earth science. *International Journal of Science Education*, *38*(11), 1699–1726. https://doi.org/10.1080/09500693.2016.1213458

Farkas, R. D. (2003). Effects of traditional versus learning-styles instructional methods on middle school students. *The Journal of Educational Research*, *97*(1), 42–51. https://doi.org/10.1080/00220670309596627

Fennema, E. H. (1972). The relative effectiveness of a symbolic and a concrete model in learning a selected mathematical principle. *Journal for Research in Mathematics Education*, *3*(4), 233–238. https://doi.org/10.5951/jresematheduc.3.4.0233

Francis, E. W. (1975). Grade Level and Task Difficulty in Learning by Discovery and Verbal Reception Methods. *Journal of Educational Psychology*, *67*(1), 146–150. https://doi.org/10.1037/h0078684

Fuchs, L. S., Fuchs, D., Prentice, K., Burch, M., Hamlett, C. L., Owen, R. Hosp, M., & Jancek, D. (2003). Explicitly teaching for transfer: Effects on third-grade students' mathematical problem solving. *Journal of Educational Psychology*, *95*(2), 293–305. https://doi.org/10.1037/0022-0663.95.2.293

Fuchs, L. S., Malone, A. S., Schumacher, R. F., Namkung, J., Hamlett, C. L., Jordan, N. C. Siegler, R. S., Gersten, R., & Changas, P. (2016). Supported self-explaining during fraction intervention. *Journal of Educational Psychology*, *108*(4), 493–508. https://doi.org/10.1037/edu0000073

Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., & Hamlett, C. L. (2010). The effects of strategic counting instruction, with and without deliberate practice, on number combination skill among students with mathematics difficulties. *Learning and Individual Differences*, *20*(2), 89–100.

https://doi.org/10.1016/j.lindif.2009.09.003

Fyfe, E. R. (2016). Providing feedback on computer-based algebra homework in middle-school classrooms. *Computers in Human Behavior*, *63*, 568–574. https://doi.org/10.1016/j.chb.2016.05.082

Fyfe, E. R., & Rittle-Johnson, B. (2016). The benefits of computer-generated feedback for mathematics problem solving. *Journal of Experimental Child Psychology*, *147*, 140–151. https://doi.org/10.1016/j.jecp.2016.03.009

Fyfe, E. R., McNeil, N. M., & Borjas, S. (2015). Benefits of "concreteness fading" for children's mathematics understanding. *Learning and Instruction*, *35*, 104–120. https://doi.org/10.1016/j.learninstruc.2014.10.004

Gellert, A. S., Arnbak, E., Wischmann, S., & Elbro, C. (2021). Morphological intervention for students with limited vocabulary knowledge: Short-and long-term transfer effects. *Reading Research Quarterly*, *56*(3), 583–601. https://doi.org/10.1002/rrq.325

Gerjets, P., Scheiter, K., & Schuh, J. (2008). Information comparisons in example-based hypermedia environments: Supporting learners with processing prompts and an interactive comparison tool. *Educational Technology Research and Development*, *56*, 73–92. https://doi.org/10.1007/s11423-007-9068-z

Glogger-Frey, I., Fleischer, C., Grüny, L., Kappich, J., & Renkl, A. (2015). Inventing a solution and studying a worked solution prepare differently for learning from direct instruction. *Learning and Instruction*, *39*, 72–87. https://doi.org/10.1016/j.learninstruc.2015.05.001

Glogger-Frey, I., Gaus, K., & Renkl, A. (2017). Learning from direct instruction: Best prepared by several self-regulated or guided invention activities?. *Learning and Instruction*, *51*, 26–35. https://doi.org/10.1016/j.learninstruc.2016.11.002

Gray, R. (2017). Transfer of training from virtual to real baseball batting. *Frontiers in Psychology*, Article 2183. https://doi.org/10.3389/fpsyg.2017.02183

Gredin, V., & Williams, A. M. (2016). The relative effectiveness of various instructional approaches during the performance and learning of motor skills. *Journal of Motor Behavior*, *48*(1), 86–97. https://doi.org/10.1080/00222895.2015.1046544

Griffin, M. M. (1995). You can′ t get there from here: Situated learning transfer, and map skills. *Contemporary Educational Psychology*, *20*(1), 65–87. https://doi.org/10.1006/ceps.1995.1004

Hoogerheide, V. (2016). *Effects of observing and creating video modeling examples on cognitive and motivational aspects of learning* (ISBN: 978-94-6299-406-5). [Doctoral dissertation, Erasmus Universiteit Rotterdam]. Ridderprint BV.

Hsu, C. Y., Kalyuga, S., & Sweller, J. (2015). When should guidance be presented in physics instruction?. *Archives of Scientific Psychology*, *3*(1), 37–53. https://doi.org/10.1037/arc0000012

Jaakkola, T., & Veermans, K. (2018). Exploring the effects of concreteness fading across grades in elementary school science education. *Instructional Science*, *46*, 185–207. https://doi.org/10.1007/s11251-017-9428-y

Jaakkola, T., & Veermans, K. (2020). Learning electric circuit principles in a simulation environment with a single representation versus "concreteness fading" through multiple representations. *Computers & Education*, *148*, Article 103811. https://doi.org/10.1016/j.compedu.2020.103811

Johnson, A. M., Reisslein, J., & Reisslein, M. (2014). Representation sequencing in computer-based engineering education. *Computers & Education*, *72*, 249–261. https://doi.org/10.1016/j.compedu.2013.11.010

Jones, C. D., & Reutzel, D. R. (2015). Write to read: Investigating the reading-writing relationship of code-level early literacy skills. *Reading & Writing Quarterly*, *31*(4), 297–315. https://doi.org/10.1080/10573569.2013.850461

Kaiser, I., Mayer, J., & Malai, D. (2018). Self-generation in the context of inquiry-based learning. *Frontiers in Psychology*, *9*, 2440. https://doi.org/10.3389/fpsyg.2018.02440

Kame‘enui, E. J., Carnine, D. W., Darch, C. B., & Stein, M. (1986). Two approaches to the development phase of mathematics instruction. *The Elementary School Journal*, *86*(5), 633–650.

Kaminski, J. A., & Sloutsky, V. M. (2020). The use and effectiveness of colorful, contextualized, student-made material for elementary mathematics instruction. *International Journal of STEM Education*, *7*(1), 1–23. https://doi.org/10.1186/s40594-019-0199-7

Kapur, M. (2010). Productive failure in mathematical problem solving. *Instructional Science*, *38*, 523–550. https://doi.org/10.1007/s11251-009-9093-x

Kapur, M. (2011). A further study of productive failure in mathematical problem solving: Unpacking the design components. *Instructional Science*, *39*, 561–579. https://doi.org/10.1007/s11251-010-9144-3

Kapur, M. (2012). Productive failure in learning the concept of variance. *Instructional Science*, *40*, 651–672. https://doi.org/10.1007/s11251-012-9209-6

Kapur, M. (2014). Comparing learning from productive failure and vicarious failure. *Journal of the Learning Sciences*, *23*(4), 651–677. https://doi.org/10.1080/10508406.2013.819000

Kapur, M. (2015). The preparatory effects of problem solving versus problem posing on learning from instruction. *Learning and Instruction*, *39*, 23–31. https://doi.org/10.1016/j.learninstruc.2015.05.004

Kapur, M. (2018). Examining the preparatory effects of problem generation and solution generation on learning from instruction. *Instructional Science*, *46*(1), 61–76. https://doi.org/10.1007/s11251-017-9435-z

Kapur, M., & Kinzer, C. K. (2009). Productive failure in CSCL groups. *International Journal of Computer-Supported Collaborative Learning*, *4*, 21–46. https://doi.org/10.1007/s11412-008-9059-z

Katzlberger, T. (2005). *Learning by teaching agents* (Publication No. 3217517). [Doctoral dissertation, Vanderbilt University]. ProQuest LLC.

Keselman, A. (2001). *Enhancing scientific reasoning by refining students' models of multivariable causality* (Publication No. 3005740). [Doctoral dissertation, Columbia University]. ProQuest LLC.

Kirschner, F., Paas, F., & Kirschner, P. A. (2009). Individual and group-based learning from complex cognitive tasks: Effects on retention and transfer efficiency. *Computers in Human Behavior*, *25*(2), 306–314. https://doi.org/10.1016/j.chb.2008.12.008

Kirschner, F., Paas, F., Kirschner, P. A., & Janssen, J. (2011). Differential effects of problem-solving demands on individual and collaborative learning outcomes. *Learning and Instruction*, *21*(4), 587–599. https://doi.org/10.1016/j.learninstruc.2011.01.001

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science*, *15*(10), 661–667. https://doi.org/10.1111/j.0956-7976.2004.00737.x

Kneppers, L., Elshout-Mohr, M., van Boxtel, C., & van Hout-Wolters, B. (2007). Conceptual learning in relation to near and far transfer in the secondary school subject of economics. *European Journal of Psychology of Education*, *22*, 115–129. https://doi.org/10.1007/BF03173517

Kramarski, B., & Mevarech, Z. R. (2003). Enhancing mathematical reasoning in the classroom: The effects of cooperative learning and metacognitive training. *American Educational Research Journal*, *40*(1), 281–310.

Kruit, P. M., Oostdam, R. J., Van den Berg, E., & Schuitema, J. A. (2018). Effects of explicit instruction on the acquisition of students' science inquiry skills in grades 5 and 6 of primary education. *International Journal of Science Education*, *40*(4), 421–441. https://doi.org/10.1080/09500693.2018.1428777

Kühl, T., & Münzer, S. (2019). The moderating role of additional information when learning with animations compared to static pictures. *Instructional Science*, *47*(6), 659–677. https://doi.org/10.1007/s11251-019-09498-x

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*, 495–503. https://doi.org/10.1007/s12564-009-9044-x

Lamnina, M., & Chase, C. C. (2021). Uncertain instruction: effects on curiosity, learning, and transfer. *Instructional Science*, *49*, 661–685. https://doi.org/10.1007/s11251-021-09557-2

Lawton, J. T., Hooper, F. H., Saunders, R. A., & Roth, P. (1984). A comparison of logical concept development in three preschool programmes. *Human Learning: Journal of Practical Research & Applications*, *3*(3), 143–164.

Lehrer, R., Lee, M., & Jeong, A. (1999). Reflective teaching of logo. *The Journal of the Learning Sciences*, *8*(2), 245–289. https://doi.org/10.1207/s15327809jls0802_3

Leopold, C., Sumfleth, E., & Leutner, D. (2013). Learning with summaries: Effects of representation mode and type of learning activity on comprehension and transfer. *Learning and Instruction*, *27*, 40–49. https://doi.org/10.1016/j.learninstruc.2013.02.003

Li, N. (2013). *Designing better scaffolding in teaching complex systems with graphical simulations* (Publication No. 3561414). [Doctoral dissertation, Columbia University]. ProQuest LLC.

Likourezos, V., & Kalyuga, S. (2017). Instruction-first and problem-solving-first approaches: Alternative pathways to learning complex tasks. *Instructional Science*, *45*, 195–219.

https://doi.org/10.1007/s11251-016-9399-4

Lindgren, R., Morphew, J. W., Kang, J., Planey, J., & Mestre, J. P. (2022). Learning and transfer effects of embodied simulations targeting crosscutting concepts in science. *Journal of Educational Psychology*, *114*(3), 462–481. https://doi.org/10.1037/edu0000697

Lonigan, C. J., Anthony, J. L., Bloomfield, B. G., Dyer, S. M., & Samwel, C. S. (1999). Effects of two shared-reading interventions on emergent literacy skills of at-risk preschoolers. *Journal of Early Intervention*, *22*(4), 306–322.

López, P., Torrance, M., Rijlaarsdam, G., & Fidalgo, R. (2021). Evaluating effects of different forms of revision instruction in upper-primary students. *Reading and Writing*, *34*(7), 1741–1767. https://doi.org/10.1007/s11145-021-10156-3

Magner, U. I., Schwonke, R., Aleven, V., Popescu, O., & Renkl, A. (2014). Triggering situational interest by decorative illustrations both fosters and hinders learning in computer-based learning environments. *Learning and Instruction*, *29*, 141–152. https://doi.org/10.1016/j.learninstruc.2012.07.002

Makransky, G., Andreasen, N. K., Baceviciute, S., & Mayer, R. E. (2021). Immersive virtual reality increases liking but not learning with a science simulation and generative learning strategies promote learning in immersive virtual reality. *Journal of Educational Psychology*, *113*(4), 719–735. https://doi.org/10.1037/edu0000473

Malkiewich, L. J., & Chase, C. C. (2019). What's your goal? The importance of shaping the goals of engineering tasks to focus learners on the underlying science. *Instructional Science*, *47*, 551–588. https://doi.org/10.1007/s11251-019-09493-2

Mason, L., Pluchino, P., & Tornatora, M. C. (2015). Eye-movement modeling of integrative reading of an illustrated text: Effects on processing and learning. *Contemporary Educational Psychology*, *41*, 172–187. https://doi.org/10.1016/j.cedpsych.2015.01.004

Mason, L., Pluchino, P., & Tornatora, M. C. (2016). Using eye-tracking technology as an indirect instruction tool to improve text and picture processing and learning. *British Journal of Educational Technology*, *47*(6), 1083–1095. https://doi.org/10.1111/bjet.12271

Matthews, P., & Rittle-Johnson, B. (2009). In pursuit of knowledge: Comparing self-explanations, concepts, and procedures as pedagogical tools. *Journal of Experimental Child Psychology*, *104*(1), 1–21. https://doi.org/10.1016/j.jecp.2008.08.004

McEldoon, K. L., Durkin, K. L., & Rittle-Johnson, B. (2013). Is self-explanation worth the time? A comparison to additional practice. *British Journal of Educational Psychology*, *83*(4), 615–632. https://doi.org/10.1111/j.2044-8279.2012.02083.x

McHugh, L., Kelly, A. M., Fisher, J. H., & David Burghardt, M. (2021). Graphing as a means to improve middle school science learning and mathematics-related affective domains. *Research in Science Education*, *51*, 301–323. https://doi.org/10.1007/s11165-018-9796-6

McNeil, N. M., Hornburg, C. B., Brletic-Shipley, H., & Matthews, J. M. (2019). Improving children’s understanding of mathematical equivalence via an intervention that goes beyond nontraditional arithmetic practice. *Journal of Educational Psychology*, *111*(6), 1023–1044. https://doi.org/10.1037/edu0000337

Melo, M., & Miranda, G. L. (2015). Learning electrical circuits: The effects of the 4C-ID instructional approach in the acquisition and transfer of knowledge. *Journal of Information Technology Education*, *14*, 313–337. https://doi.org/10.28945/2281

Meneses, A., Escobar, J. P., & Véliz, S. (2018). The effects of multimodal texts on science reading comprehension in Chilean fifth-graders: text scaffolding and comprehension skills. *International Journal of Science Education*, *40*(18), 2226–2244. https://doi.org/10.1080/09500693.2018.1527472

Moore, N. S., & MacArthur, C. A. (2012). The effects of being a reader and of observing readers on fifth-grade students' argumentative writing and revising. *Reading and Writing*, *25*, 1449–1478. https://doi.org/10.1007/s11145-011-9327-6

Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents?. *Cognition and Instruction*, *19*(2), 177–213. https://doi.org/10.1207/S1532690XCI1902_02

Moreno, R., Ozogul, G., & Reisslein, M. (2011). Teaching with concrete and abstract visual representations: Effects on students' problem solving, problem representations, and learning perceptions. *Journal of Educational Psychology*, *103*(1), 32–47. https://doi.org/10.1037/a0021995

Mullins, D., Rummel, N., & Spada, H. (2011). Are two heads always better than one? Differential effects of collaboration on students' computer-supported learning in mathematics. *International Journal of Computer-Supported Collaborative Learning*, *6*, 421–443. https://doi.org/10.1007/s11412-011-9122-z

Muthukrishna, N., & Borkowski, J. G. (1995). How learning contexts facilitate strategy transfer. *Applied Cognitive Psychology*, *9*(5), 425–446. https://doi.org/10.1002/acp.2350090506

Ngu, B. H., Mit, E., Shahbodin, F., & Tuovinen, J. (2009). Chemistry problem solving instruction: a comparison of three computer-based formats for learning from hierarchical network problem representations. *Instructional Science*, *37*, 21–42. https://doi.org/10.1007/s11251-008-9072-7

Ngu, B. H., Yeung, A. S., & Phan, H. P. (2015). Constructing a coherent problem model to facilitate algebra problem solving in a chemistry context. *International Journal of Mathematical Education in Science and Technology*, *46*(3), 388–403. https://doi.org/10.1080/0020739X.2014.979899

Nunes, T., Bryant, P., Hallett, D., Bell, D., & Evans, D. (2009). Teaching children about the inverse relation between addition and subtraction. *Mathematical Thinking and Learning*, *11*(1-2), 61–78. https://doi.org/10.1080/10986060802583980

Obiagu, A. N., Mezieobi, D. I., Aroh, P. N., & Akubue, F. N. (2020). The effect of cooperative concept mapping on misconceptions, knowledge achievement, and transfer of learning in peace education. *The Social Studies*, *111*(1), 18–38. https://doi.org/10.1080/00377996.2019.1652139

Oostdam, R., Blok, H., & Boendermaker, C. (2015). Effects of individualised and small-group guided oral reading interventions on reading skills and reading attitude of poor readers in grades 2-4. *Research Papers in Education*, *30*(4), 427–450. https://doi.org/10.1080/02671522.2014.953195

Paek, S. (2012). *The impact of multimodal virtual manipulatives on young children's mathematics learning* (Publication No. 3554708). [Doctoral dissertation, Teachers College, Columbia University]. ProQuest LLC.

Papaevripidou, M., Constantinou, C. P., & Zacharia, Z. C. (2007). Modeling complex marine ecosystems: An investigation of two teaching approaches with fifth graders. *Journal of Computer Assisted Learning*, *23*(2), 145–157. https://doi.org/10.1111/j.1365-2729.2006.00217.x

Petersen, G. B., Klingenberg, S., Mayer, R. E., & Makransky, G. (2020). The virtual field trip: Investigating how to optimize immersive virtual learning in climate change education. *British Journal of Educational Technology*, *51*(6), 2099–2115. https://doi.org/10.1111/bjet.12991

Philippakos, Z. A., & MacArthur, C. A. (2016). The effects of giving feedback on the persuasive writing of fourth-and fifth-grade students. *Reading Research Quarterly*, *51*(4), 419–433. https://doi.org/10.1002/rrq.149

Pritchard, T., Hawkins, A., Wiegand, R., & Metzler, J. N. (2008). Effects of two instructional approaches on skill development, knowledge, and game performance. *Measurement in Physical Education and Exercise Science*, *12*(4), 219–236. https://doi.org/10.1080/10913670802349774

Purnama, P. W., & Retnowati, E. (2020). The effectiveness of goal-free problems for studying triangle similarity in collaborative groups. *Journal of Research and Advances in Mathematics Education*, *6*(1), 32–45. https://doi.org/10.23917/jramathedu.v6i1.11198

Rau, M. A., Aleven, V., & Rummel, N. (2015). Successful learning with multiple graphical representations and self-explanation prompts. *Journal of Educational Psychology*, *107*(1), 30–46. https://doi.org/10.1037/a0037211

Reitsma, P., & Wesseling, R. (1998). Effects of computer-assisted training of blending skills in kindergartners. *Scientific Studies of Reading*, *2*(4), 301–320. https://doi.org/10.1207/s1532799xssr0204_1

Ren, K., & Gunderson, E. A. (2021). The dynamic nature of children's strategy use after receiving accuracy feedback in decimal comparisons. *Journal of Experimental Child Psychology*, *202*, 105015. https://doi.org/10.1016/j.jecp.2020.105015

Renkl, A., Atkinson, R. K., Maier, U. H., & Staley, R. (2002). From example study to problem solving: Smooth transitions help learning. *The Journal of Experimental Education*, *70*(4), 293–315. https://doi.org/10.1080/00220970209599510

Rittle-Johnson, B. (2006). Promoting transfer: Effects of self-explanation and direct instruction. *Child Development*, *77*(1), 1–15. https://doi.org/10.1111/j.1467-8624.2006.00852.x

Rittle-Johnson, B., Saylor, M., & Swygert, K. E. (2008). Learning from explaining: Does it matter if mom is listening?. *Journal of Experimental Child Psychology*, *100*(3), 215–224. https://doi.org/10.1016/j.jecp.2007.10.002

Ross, L. (1991). *Mediating science learning: promoting the acquisition and transfer of biological concepts* (Publication No. 9210969). [Doctoral dissertation, University of

Illinois at Urbana-Champaign.] ProQuest LLC.

Safadi, R., & Yerushalmi, E. (2014). Problem solving vs. troubleshooting tasks: The case of sixth-grade students studying simple electric circuits. *International Journal of Science and Mathematics Education*, *12*, 1341–1366. https://doi.org/10.1007/s10763-013-9461-5

Schalk, L., Schumacher, R., Barth, A., & Stern, E. (2018). When problem-solving followed by instruction is superior to the traditional tell-and-practice sequence. *Journal of Educational Psychology*, *110*(4), 596–610. https://doi.org/10.1037/edu0000234

Scheiter, K., Gerjets, P., & Schuh, J. (2010). The acquisition of problem-solving skills in mathematics: How animations can aid understanding of structural problem features and solution procedures. *Instructional Science*, *38*, 487–502. https://doi.org/10.1007/s11251-009-9114-9

Scheiter, K., Schüler, A., Gerjets, P., Huk, T., & Hesse, F. W. (2014). Extending multimedia research: How do prerequisite knowledge and reading comprehension affect learning from text and pictures. *Computers in Human Behavior*, *31*, 73–84. https://doi.org/10.1016/j.chb.2013.09.022

Sidney, P. G., & Alibali, M. W. (2015). Making connections in math: Activating a prior knowledge analogue matters for learning. *Journal of Cognition and Development*, *16*(1), 160–185. https://doi.org/10.1080/15248372.2013.792091

Smith, J. G., & Smith, R. L. (2012). Screen-capture instructional technology: A cognitive tool for designing a blended multimedia curriculum. *Journal of Educational Computing Research*, *46*(3), 207–228. https://doi.org/10.2190/EC.46.3.a

Star, J. R., & Rittle-Johnson, B. (2008). Flexibility in problem solving: The case of equation solving. *Learning and Instruction*, *18*(6), 565–579. https://doi.org/10.1016/j.learninstruc.2007.09.018

Star, J. R., & Rittle-Johnson, B. (2009). It pays to compare: An experimental study on computational estimation. *Journal of Experimental Child Psychology*, *102*(4), 408–426. https://doi.org/10.1016/j.jecp.2008.11.004

Tajika, H., Nakatsu, N., Nozaki, H., Neumann, E., & Maruno, S. (2007). Effects of self-explanation as a metacognitive strategy for solving mathematical word problems. *Japanese Psychological Research*, *49*(3), 222–233. https://doi.org/10.1111/j.1468-5884.2007.00349.x

Tang, M., Ginns, P., & Jacobson, M. J. (2019). Tracing enhances recall and transfer of knowledge of the water cycle. *Educational Psychology Review*, *31*, 439–455. https://doi.org/10.1007/s10648-019-09466-4

Terwel, J., van Oers, B., van Dijk, I., & van den Eeden, P. (2009). Are representations to be provided or generated in primary mathematics education? Effects on transfer. *Educational Research and Evaluation*, *15*(1), 25–44. https://doi.org/10.1080/13803610802481265

Travlos, A. K. (2010). Specificity and variability of practice, and contextual interference in acquisition and transfer of an underhand volleyball serve. *Perceptual and Motor Skills*, *110*(1), 298–312. https://doi.org/10.2466/pms.110.1.298-312

Vadasy, P. F., Sanders, E. A., & Nelson, J. R. (2015). Effectiveness of supplemental kindergarten vocabulary instruction for English learners: A randomized study of immediate and longer-term effects of two approaches. *Journal of Research on Educational Effectiveness*, *8*(4), 490–529. https://doi.org/10.1080/19345747.2015.1033120

Van Eck, R., & Dempsey, J. (2002). The effect of competition and contextualized advisement on the transfer of mathematics skills a computer-based instructional simulation game. *Educational Technology Research and Development*, *50*(3), 23–41. https://doi.org/10.1007/BF02505023

Van Gog, T., Paas, F., & Van Merriënboer, J. J. G. (2008). Effects of studying sequences of process-oriented and product-oriented worked examples on troubleshooting transfer efficiency. *Learning and Instruction*, *18*, 211–222. https://doi.org/10.1016/j.learninstruc.2007.03.003

Van Meter, P., Aleksic, M., Schwartz, A., & Garner, J. (2006). Learner-generated drawing as a strategy for learning from content area text. *Contemporary Educational Psychology*, *31*(2), 142–166. https://doi.org/10.1016/j.cedpsych.2005.04.001

Wagensveld, B., Segers, E., Kleemans, T., & Verhoeven, L. (2015). Child predictors of learning to control variables via instruction or self-discovery. *Instructional Science*, *43*, 365–379. https://doi.org/10.1007/s11251-014-9334-5

Wang, Z., Ardasheva, Y., Carbonneau, K., & Liu, Q. (2021). Testing the seductive details effect: Does the format or the amount of seductive details matter?. *Applied Cognitive Psychology*, *35*(3), 761–774. https://doi.org/10.1002/acp.3801

Wideman, H. H., & Owston, R. D. (1993). Knowledge base construction as a pedagogical activity. *Journal of Educational Computing Research*, *9*(2), 165–196. https://doi.org/10.2190/LUHJ-5MRH-WT8T-1J6N

Williams, P. B., & Carnine, D. W. (1981). Relationship between range of examples and of instructions and attention in concept attainment. *The Journal of Educational Research*, *74*(3), 144–148. https://doi.org/10.1080/00220671.1981.10885300

Wong, R. M., Lawson, M. J., & Keeves, J. (2002). The effects of self-explanation training on students' problem solving in high-school mathematics. *Learning and Instruction*, *12*(2), 233–262. https://doi.org/10.1016/S0959-4752(01)00027-5

Wouters, P., Paas, F., & van Merriënboer, J. J. (2010). Observational learning from animated models: Effects of studying-practicing alternation and illusion of control on transfer. *Instructional Science*, *38*(1), 89–104. https://doi.org/10.1007/s11251-008-9079-0

Wulf, G., Chiviacowsky, S., & Cardozo, P. L. (2014). Additive benefits of autonomy support and enhanced expectancies for motor learning. *Human Movement Science*, *37*, 12–20. https://doi.org/10.1016/j.humov.2014.06.004

Yang, Z. K., Wang, M., Cheng, H. N., Liu, S. Y., Liu, L., & Chan, T. W. (2016). The Effects of learning from correct and erroneous examples in individual and collaborative settings. *The Asia-Pacific Education Researcher*, *25*, 219–227. https://doi.org/10.1007/s40299-015-0253-2

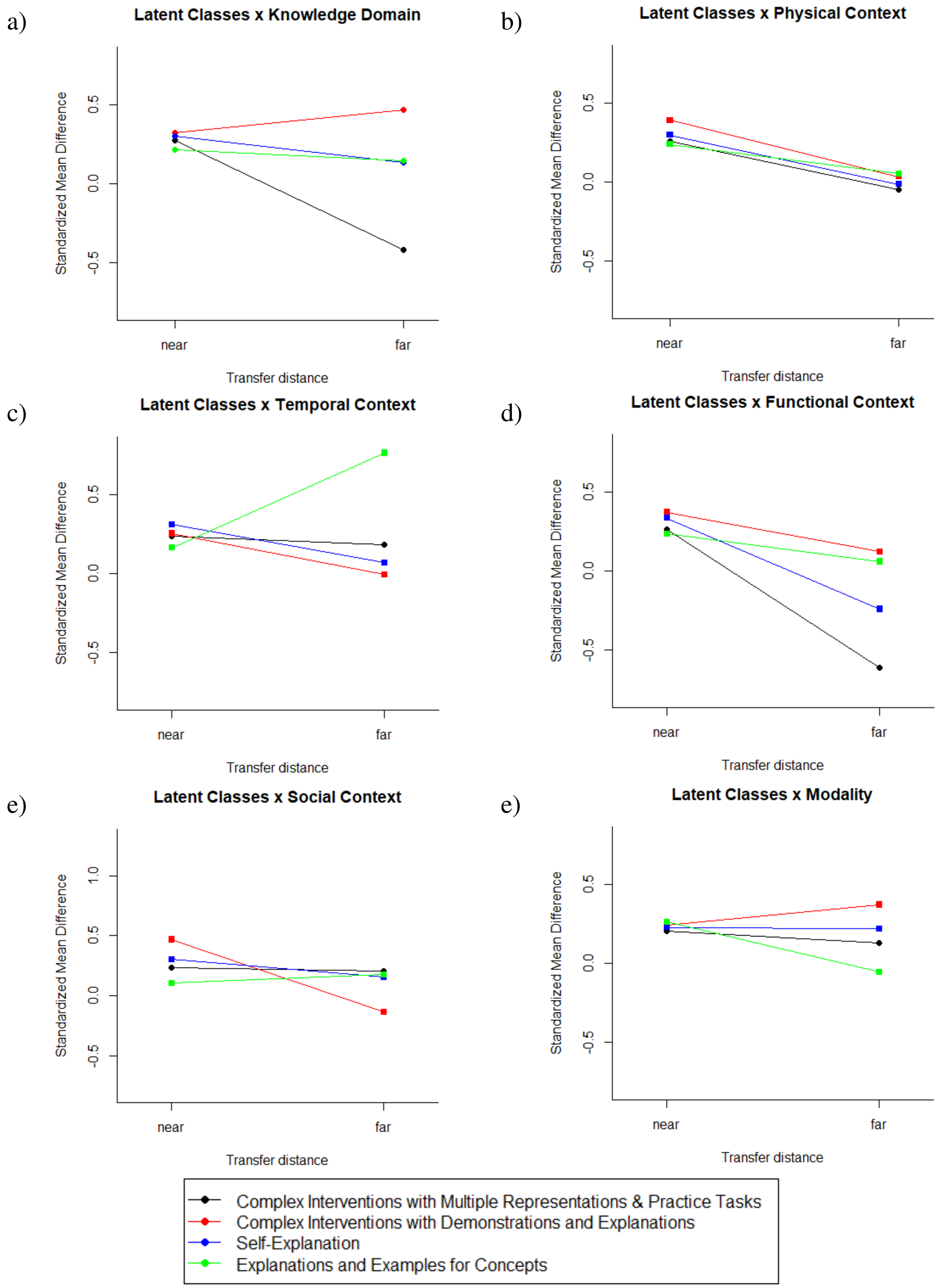
Zhang, L. (2019). "Hands-on" plus "inquiry"? Effects of withholding answers coupled with physical manipulations on students' learning of energy-related science concepts. *Learning and Instruction*, *60*, 199–205. https://doi.org/10.1016/j.learninstruc.2018.01.001

Zhang, L., Shang, J., Pelton, T., & Pelton, L. F. (2020). Supporting primary students' learning of fraction conceptual knowledge through digital games. *Journal of Computer Assisted Learning*, *36*(4), 540–548. https://doi.org/10.1111/jcal.12422

Appendix I: Interaction Latent Classes and Transfer Distance (Study 3)

Figure 48

Interaction Latent Classes and Transfer Distance on Context Factors



Note. Interaction of the four latent classes and the transfer distance on the context factors a) knowledge domain, b) physical context, c) temporal context, d) functional context, e) social context, and f) modality. On the y-axis is the mean effect as standardized mean difference (Hedges' *g*).

Appendix J: Mean Effect Sizes and Moderator Analyses (Study 3)

Table 22

Mean Effect Sizes and Moderator Analyses for the Comparison of Transfer Knowledge of Experimental and Comparison Groups

	Subgroup analysis								Moderator analysis		
	j	*k*	g^+	CI g^+ 95%	*Sign.*	σ^1	σ^2	I^2	g^+_{MOD}	*Sign.*	R^2
Overall	151	797	0.285	[.175; .395]	***	.386	.301	91.05	-	-	-
Intervention characteristics											
Setting Intervention										*	.000
Field	131	665	0.312	[.189; .436]	***	.475	.314	93.07	0.311	*	
Lab	7	27	-0.143	[-.474; .187]	*ns*	.085	.109	71.64	-0.149	Ref.	
Mean number of sessions	146	789								*ns*	.000
Mean duration of sessions (minutes)	120	645								*ns*	.060
Mean days between intervention and transfer test	130	740								*ns*	.075
Broad content domain										*ns*	.000
Language	13	49	0.243	[-.077; .562]	*ns*	.245	.225	92.37	0.243	Ref.	
Movement	7	98	-0.001	[-.946; .945]	*ns*	4.228	.846	97.09	-0.031	*ns*	
Other	1	2	1.254	[-2.759; 5.267]	*ns*	.000	.000	00.00	-	-	
Social Sciences	1	4	0.202	[-.633; 1.036]	*ns*	.142	.000	51.56	-	-	
STEM	128	643	0.300	[.188; .411]	***	.243	.283	88.92	0.301	*ns*	
Content domain										*ns*	.000
Biology	17	95	0.486	[.060; .912]	*	.192	.673	90.90	0.498	*ns*	
Chemistry	4	19	0.126	[-.392; .645]	*ns*	.139	.171	72.56	-	-	

Table 22 (continued)

	Subgroup analysis								Moderator analysis		
	j	*k*	g^+	CI g^+ 95%	*Sign.*	σ^1	σ^2	I^2	g^+_{MOD}	*Sign.*	R^2
Engineering	4	32	0.159	[-.090; .408]	*ns*	.078	.032	66.60	-	-	
First language	10	40	0.206	[-.199; .611]	*ns*	.259	.292	91.98	0.208	*ns*	
Foreign language	3	9	0.410	[.043; .776]	*	.191	.000	89.98	-	-	
Geosciences	6	19	0.092	[-.160; .349]	*ns*	.192	.000	69.44	0.134	Ref.	
Mathematics	70	298	0.314	[.176; .451]	***	.264	.199	88.10	0.322	*ns*	
Other	6	30	0.510	[-.072; 1.091]	*ns*	.081	.424	94.55	0.494	*ns*	
Physics	20	145	0.149	[-.147; .445]	*ns*	.355	.326	89.70	0.148	*ns*	
Several	2	9	0.036	[-.553; .625]	*ns*	.070	.131	63.44	-	-	
Sports	7	98	-0.001	[-.946; .945]	*ns*	4.228	.846	97.09	-0.033	*ns*	
Group size of intervention										*ns*	.111
Single	38	187	0.358	[.179; .536]	***	.074	.220	73.70	0.367	Ref.	
Group	96	460	0.248	[.096; .399]	**	.313	.420	93.23	0.247	*ns*	
Several	4	9	0.378	[-.039; .795]	*ns*	.000	.101	61.25	-	-	
Trainer person										*ns*	.000
Mixed	7	54	-0.241	[-1.157; .674]	*ns*	.061	1.222	95.16	-0.163	Ref.	
Other	2	7	0.343	[-.232; .918]	*ns*	.307	.000	80.81	-	-	
Researcher	69	336	0.279	[.093; .465]	**	.417	.433	91.93	0.276	*ns*	
Teacher	32	123	0.362	[.157; .567]	***	.101	.262	88.73	0.395	*ns*	
Technology-enhanced learning										*ns*	.000
Yes	62	367	0.341	[.180; .501]	***	.220	.308	85.92	0.245	*ns*	
No	88	426	0.243	[.088; .397]	**	.698	.262	94.39	0.341	Ref.	
Intelligent tutor										-	-
Yes	3	49	0.135	[-.022; .292]	*ns*	.223	.000	76.34	-	-	
No	148	723	0.282	[.159; .405]	***	.441	.333	91.57	-	-	

Table 22 (continued)

	Subgroup analysis								Moderator analysis		
	j	*k*	g^+	CI g^+ 95%	*Sign.*	σ^1	σ^2	I^2	g^+_{MOD}	*Sign.*	R^2
Intention to transfer										-	-
Yes	3	10	0.926	[.666; 1.186]	**	.000	.000	00.00	-	-	
No	148	784	0.276	[.161; .391]	***	.400	.311	90.76	-	-	
Instructional methods in intervention											
Number of methods	151	797								*ns*	.000
Multivariate effect of all instructional methods	151	797								*ns*	.002
Advance Organizer										-	-
Yes	2	13	0.421	[.236; .606]	***	.036	.000	40.30	-	-	
No	149	780	0.285	[.172; 398]	***	.398	.311	91.16	-	-	
Annotated Examples										*ns*	.000
Yes	28	76	0.368	[.163; 572]	***	.100	.209	79.89	0.397	*ns*	
No	123	721	0.264	[.143; .384]	***	.433	.311	91.78	0.265	Ref.	
Collaborative Learning										*ns*	.000
Yes	10	32	0.289	[-.040; .617]	*ns*	.222	.133	91.51	0.323	*ns*	
No	141	765	0.286	[.170; .401]	***	.397	.313	90.84	0.283	Ref.	
Comparison of Annotated Examples										*ns*	.000
Yes	9	27	0.160	[-.232; .551]	*ns*	.000	.287	78.07	0.229	*ns*	
No	142	770	0.292	[.178; .406]	***	.405	.306	91.37	0.288	Ref.	
Comparison of Concepts										*ns*	.000
Yes	8	58	0.366	[-.014; .747]	*ns*	.211	.263	92.60	0.413	*ns*	
No	143	742	0.281	[.168; 394]	***	.406	.299	90.75	0.276	Ref.	

Table 22 (continued)

	Subgroup analysis								Moderator analysis		
	j	*k*	g^+	CI g^+ 95%	*Sign.*	σ^1	σ^2	I^2	g^+_{MOD}	*Sign.*	R^2
Concrete Models										*ns*	.006
Yes	11	51	0.556	[.003; 1.109]	*	.044	.931	92.67	0.476	*ns*	
No	140	746	0.252	[.150; .355]	***	.407	.214	90.26	0.268	Ref.	
Concreteness Fading										*ns*	.003
Yes	7	33	0.103	[-.408; .614]	*ns*	.269	.577	93.38	0.148	*ns*	
No	144	764	0.296	[.188; .405]	***	.394	.269	90.70	0.293	Ref.	
Demonstration										*ns*	.004
Yes	21	89	0.456	[.401; .871]	*	.233	.798	92.63	0.365	*ns*	
No	130	708	0.243	[.141; .345]	***	.416	.194	90.23	0.273	Ref.	
Drill										*ns*	.000
Yes	6	59	0.235	[-.377; .846]	*ns*	.268	.457	91.90	0.301	*ns*	
No	145	738	0.289	[.178; .400]	***	.407	.289	91.12	0.284	Ref.	
Examples of Concepts										*ns*	.015
Yes	19	100	0.567	[.088; 1.046]	*	.710	.809	97.02	0.540	*ns*	
No	132	697	0.240	[.136; .344]	***	.337	.225	88.51	0.250	Ref.	
Explanations										*ns*	.000
Yes	25	120	0.232	[.050; .415]	*	.441	.063	87.12	0.199	*ns*	
No	126	677	0.291	[.170; .412]	***	.373	.330	91.36	0.299	Ref.	
Feedback										*ns*	.003
Yes	18	84	0.116	[-.302; .533]	*ns*	.350	.746	94.81	0.120	*ns*	
No	133	716	0.306	[.201; .412]	***	.392	.223	90.11	0.309	Ref.	
Generate Drawings										-	-
Yes	4	32	0.936	[-.427; 2.300]	*ns*	.760	2.445	96.62	-	-	

Table 22 (continued)

	Subgroup analysis								Moderator analysis		
	j	k	g^+	CI g^+ 95%	*Sign.*	σ^1	σ^2	I^2	g^+_{MOD}	*Sign.*	R^2
No	147	765	0.254	[.156; .352]	***	.381	.204	89.79	-	-	
Imitation										*ns*	.000
Yes	5	28	0.132	[-.134; .398]	*ns*	.005	.067	47.74	0.001	*ns*	
No	146	769	0.292	[.178; .405]	***	.407	.310	91.42	0.295	Ref.	
Interleaved Practice										-	-
Yes	1	4	0.847	[-.499; 2.194]	*ns*	.546	.000	76.53	-	-	
No	150	793	0.284	[.174; .395]	***	.386	.301	91.07	-	-	
Mnemonic Systems										-	-
Yes	1	4	-0.250	[-.621; .122]	*ns*	.000	.000	00.00	-	-	
No	150	793	0.289	[.178; .400]	***	.390	.302	91.09	-	-	
Movement										*ns*	.000
Yes	10	34	0.378	[.125; .630]	**	.120	.091	71.56	0.421	*ns*	
No	141	763	0.276	[.161; 391]	***	.404	.314	91.47	0.276	Ref.	
Multiple Representations										*ns*	.008
Yes	22	186	0.543	[.262; .823]	***	.352	.438	92.31	0.384	*ns*	
No	129	607	0.227	[.115; .339]	***	.413	.240	90.36	0.263	Ref.	
Paraphrasing										-	-
Yes	1	1	-	-	-	-	-	-	-	-	
No	150	796	0.289	[.177; .397]	***	.385	.304	91.06	-	-	
Problem Posing										-	-
Yes	4	10	0.308	[-.252; .868]	*ns*	.233	.094	83.85	-	-	
No	147	787	0.285	[.172; .397]	***	.389	.308	91.16	-	-	

Table 22 (continued)

	Subgroup analysis								Moderator analysis		
	j	*k*	g^+	CI g^+ 95%	*Sign.*	σ^1	σ^2	I^2	g^+_{MOD}	*Sign.*	R^2
Realistic Problems										*ns*	.009
Yes	9	30	0.666	[.134; 1.200]	*	.388	.448	96.66	0.550	*ns*	
No	142	767	0.257	[.146; .367]	***	.387	.279	90.21	0.269	Ref.	
Rehearsal										-	-
Yes	2	6	0.117	[-.175; .409]	*ns*	.000	.011	20.23	-	-	
No	149	791	0.290	[.176; .399]	***	.391	.306	91.13	-	-	
Scaffolding										*ns*	.000
Yes	14	63	0.148	[-.078; .373]	*ns*	.077	.133	82.68	0.160	*ns*	
No	137	734	0.299	[.178; .419]	***	.437	.322	91.46	0.299	Ref.	
Self-Explanations										*ns*	.000
Yes	21	94	0.289	[.092; .495]	**	.047	.182	76.90	0.320	*ns*	
No	130	703	0.283	[.161; .405]	***	.470	.317	92.10	0.280	Ref.	
Variability of Practice										*ns*	.000
Yes	25	156	0.312	[-.041; .665]	*ns*	2.376	.212	97.49	0.507	*ns*	
No	126	644	0.265	[.157; 374]	***	.255	.274	88.71	0.244	Ref.	
Verbal or Written Instructions										*ns*	.001
Yes	10	63	0.032	[-.602; .665]	*ns*	.044	.947	91.00	0.190	*ns*	
No	141	734	0.298	[.193; .404]	***	.429	.236	90.99	0.291	Ref.	
Visualizations										*ns*	.000
Yes	16	67	0.347	[-.078; .772]	*ns*	.155	.757	91.36	0.369	*ns*	
No	135	730	0.267	[.162; .373]	***	.409	.225	90.54	0.275	Ref.	
Latent Classes											.000
Multiple Representations & Variability	99	530	0.214	[.084; .343]	**	.485	.237	91.75	0.260	*ns*	

Table 22 (continued)

	Subgroup analysis								Moderator analysis		
	j	*k*	g^+	CI g^+ 95%	*Sign.*	σ^1	σ^2	I^2	g^+_{MOD}	*Sign.*	R^2
Demonstration	25	86	0.448	[.083; .814]	*	.253	.715	92.61	0.373	*ns*	
Self-Explanations	31	115	0.315	[.156; .472]	***	.059	.135	72.58	0.327	*ns*	
Examples & Explanations	12	66	0.258	[-.014; .529]	*ns*	.753	.026	91.60	0.233	Ref.	
Learner characteristics experimental group											
Age	85	428								*ns*	.000
Age group										*	.000
Under 6 years	10	47	0.646	[.260; 1.031]	**	.140	.231	88.42	0.635	Ref.	
6 to 11 years	38	164	0.145	[-.138; .427]	*ns*	2.387	.090	96.87	0.150	*ns*	
12 to 18 years	61	279	0.154	[.027; .280]	*	.378	.108	88.43	0.152	*	
Several	1	20	0.102	[-.096; .300]	*ns*	.105	.000	58.84	-	-	
Educational level										*	.000
Kindergarten/preschool	13	83	0.836	[.384; 1.288]	***	3.198	.088	98.19	0.919	Ref.	
Primary school	51	251	0.290	[.049; .531]	*	.335	.613	93.28	0.152	*	
Secondary school	82	437	0.216	[.106; .326]	***	.275	.143	86.01	0.154	*	
Several	4	13	-0.083	[-.343; .177]	*ns*	.087	.000	47.56	-	-	
Class	128	656								*	.000
Gender										-	-
Mixed	117	530	0.266	[.178; .354]	***	.264	.111	86.66	-	-	
Female	1	36	0.263	[-2.729; 3.254]	*ns*	2.157	6.218	98.09	-	-	
Male	4	77	0.433	[-1.088; 1.953]	*ns*	1.987	2.900	97.24	-	-	
Learner characteristics comparison group											
Age	85	428								*	.000

Table 22 (continued)

	Subgroup analysis								Moderator analysis		
	j	*k*	g^+	CI g^+ 95%	*Sign.*	σ^1	σ^2	I^2	g^+_{MOD}	*Sign.*	R^2
Age group										*	.000
Under 6 years	10	47	0.646	[.260; 1.031]	**	.140	.231	88.42	0.635	Ref.	
6 to 11 years	38	163	0.141	[-.142; .425]	*ns*	2.414	.086	96.90	0.152	*ns*	
12 to 18 years	61	278	0.154	[.026; .282]	*	.379	.112	88.47	0.149	*	
Several	1	20	0.102	[-.096; .300]	*ns*	.105	.000	58.84	-	-	
Educational level										**	.000
Kindergarten/preschool	13	83	0.836	[.383; 1.288]	***	3.198	.088	98.19	0.944	Ref.	
Primary school	51	251	0.304	[.061; .548]	*	.360	.618	93.47	0.256	*	
Secondary school	80	437	0.202	[.098; .307]	***	.285	.117	85.55	0.202	**	
Several	4	13	-0.083	[-.343; .177]	*ns*	.087	.000	47.56	-	-	
Class	128	656								**	.000
Gender										-	-
Mixed	117	531	0.267	[.172; .362]	***	.285	.120	87.83	-	-	
Female	1	36	1.705	[.359; 3.051]	*	7.900	.000	96.92	-	-	
Male	4	77	0.000	[-1.725; 1.725]	*ns*	2.440	3.767	97.78	-	-	
Transfer characteristics											
Theoretical basis										*ns*	.000
Abstraction	4	13	0.479	[-.366; 1.325]	*ns*	.635	.334	88.78	-	-	
Analogical Transfer	4	28	-0.454	[-2.166; 1.258]	*ns*	.201	2.047	95.60	-	-	
Formal Discipline	1	9	0.255	[.003; .506]	*	.000	.000	00.00	-	-	
Identical Elements	4	17	0.278	[.059; .497]	*	.069	.007	51.92	-	-	
Other	2	80	0.482	[-.049; 1.013]	*ns*	5.168	.007	97.20	-	-	

Table 22 (continued)

	Subgroup analysis								Moderator analysis		
	j	*k*	g^+	CI g^+ 95%	*Sign.*	σ^1	σ^2	I^2	g^+_{MOD}	*Sign.*	R^2
Preparation for Future Learning	1	1	-	-	-	-	-	-	-	-	
Several	6	17	0.743	[.433; 1.052]	***	.011	.101	73.99	0.744	*ns*	
Situated Learning	5	13	0.172	[-.193; .537]	*ns*	.000	.097	48.23	0.170	*Ref.*	
Knowledge domain										*ns*	.000
Near	138	745	0.279	[.164; .393]	***	.407	.292	91.31	0.283	*ns*	
Far	9	44	0.400	[-.039; .838]	*ns*	.127	.609	89.35	0.261	Ref.	
Physical context										-	-
Near	140	703	0.285	[.166; .404]	***	.443	.319	92.04	-	-	
Far	4	13	0.268	[.074; .461]	*	.000	.001	1.36	-	-	
Temporal context										*ns*	.270
Near	117	563	0.249	[.139; .359]	***	.260	.231	87.14	0.252	*ns*	
Far	12	146	0.403	[.158; .647]	**	.182	.308	88.80	0.373	Ref.	
Functional context										-	-
Near	149	790	.290	[.178; .402]	***	.391	.305	91.13	-	-	
Far	2	7	.018	[-488; .524]	*ns*	.039	.085	72.32	-	-	
Social context										*ns*	.139
Near	105	505	0.249	[.120; .378]	***	.259	.312	89.89	0.267	*ns*	
Far	11	43	0.473	[.011; .935]	*	.379	.461	93.15	0.419	Ref.	
Modality										*ns*	.000
Near	100	481	0.198	[.059; .337]	**	.703	.232	93.44	0.221	*ns*	
Far	47	290	0.427	[.246; .609]	***	.111	.364	86.64	0.395	Ref.	
Multivariate effect of all transfer dimensions	93	475								*ns*	.261

Table 22 (continued)

	Subgroup analysis								Moderator analysis		
	j	*k*	g^+	CI g^+ 95%	*Sign.*	σ^1	σ^2	I^2	g^+_{MOD}	*Sign.*	R^2
Methodological study characteristics											
Grey literature										*ns*	.014
No	137	721	0.337	[.082; .592]	*	.602	.058	91.02	0.268	*ns*	
Yes	12	68	0.270	[.151; .388]	***	.338	.333	88.96	0.375	Ref.	
Study design										*ns*	.000
RCT	114	657	0.236	[.133; .339]	***	.442	.150	90.05	0.249	*ns*	
Quasi-experimental	31	121	0.337	[-.011; .684]	*ns*	.210	.790	92.50	0.369	Ref.	
Longitudinal design										*ns*	.000
Pre-post	39	201	0.282	[.055; .509]	*	.384	.362	93.40	0.286	Ref.	
Post only	110	581	0.281	[.152; .411]	***	.413	.297	90.51	0.294	*ns*	
Setting pre-test										-	-
Lab	2	5	0.038	[-.503; .578]	*ns*	.046	.000	24.22	-	-	
Field	37	188	0.288	[.031; .545]	*	.407	.420	94.09	-	-	
Setting post-test										*ns*	.000
Lab	6	25	-0.012	[-.506; .482]	*ns*	.094	.195	72.92	-0.117	Ref.	
Field	132	667	0.310	[.181; .438]	***	.504	.328	93.07	0.318	*ns*	
Number of items in transfer knowledge measure	133	640								*ns*	.000
Response format										*ns*	.293
Mixed	12	31	0.393	[.174; .613]	***	.031	.107	80.22	0.337	Ref.	
Multiple choice	20	91	0.147	[-.055; .350]	*ns*	.092	.177	51.22	0.234	*ns*	
Open	101	544	0.334	[.212; .455]	***	.279	.275	90.60	0.334	*ns*	

Table 22 (continued)

	Subgroup analysis								Moderator analysis		
	j	k	g^+	CI g^+ 95%	*Sign.*	σ^1	σ^2	I^2	g^+_{MOD}	*Sign.*	R^2
Implementation check										**	.031
Yes, successful	37	188	0.602	[.438; .767]	***	.178	.173	85.04	0.772	Ref.	
Yes, not successful	67	303	0.082	[-.042; .205]	*ns*	.255	.160	86.33	0.012	**	
No	45	284	0.339	[.046; .633]	*	1.060	.629	95.73	0.319	*	
Continent										*ns*	.012
Africa	1	18	-0.071	[-.291; .149]	*ns*	.117	.000	59.64	-	-	
Asia	25	102	0.509	[.120; .898]	*	.714	.608	94.66	0.497	*ns*	
Australia	8	11	0.442	[-.011; .896]	*ns*	.000	.251	72.07	0.400	Ref.	
Europe	40	226	0.097	[-.148; .342]	*ns*	.289	.480	91.16	0.082	*ns*	
North America	75	438	0.310	[.205; .436]	***	.416	.118	89.45	0.325	*ns*	
South America	2	2	0.310	[-6.179; 6.799]	*ns*	.247	.247	94.85	-	-	

Note. j - Number of studies, k - Number of effect sizes, g^+ - Hedges' g (subgroup), 95% CI - 95% confidence interval, Sign – Significance (* $p < .05$; ** $p < .01$; *** < .001, *ns* – nonsignificant); σ^1 - Within-study variance, σ^2 - Between-study variance, I^2 - Indicator for heterogeneity, g^+_{MOD} - Hedges' g (moderator), Ref - Indicates the reference category used for the respective categorical moderator analysis; each other moderator level was tested against this reference category. R^2 - Variance explanation. Missing values are indicated by -. This means that the respective analysis was not possible because the number of effect sizes on study level for this analysis was too small.

Appendix K: Risk of Bias Assessment of Included Studies (Study 3)

Table 23

Questions for Risk of Bias Assessment of Included Studies (Study 3)

1	**Baseline Confounding**	**Yes**	**No**	**Not applicable**
1a	Potential baseline confounding was addressed (e.g., by design or adjustments).			
	Rating:			
2	**Sample selection bias**	**Yes**	**No**	**Not applicable**
2a	The drop-out rate is documented and presented.			
2b	The drop-out rate is below 20%.			
2c	Exact drop-out in numbers:			
	Rating:			
3	**Bias in measurement**	**Yes**	**No**	**Not applicable**
3a	The method of measurement was appropriate.			
3b	The reliability of all measurements is mentioned.			
3c	The reported reliability is adequately high.			
	Rating:			
4	**Selective reporting of outcomes**	**Yes**	**No**	**Not applicable**
4a	The numerical result is not likely to have been selected from multiple measurements.			
	Rating:			
5	**Study design**	**Yes**	**No**	**Not applicable**
5a	The data includes pre- and post-measurements of knowledge transfer.			
	Rating:			
6	**Funding**	**Yes**	**No**	**Not applicable**
6a	The study was not externally and commercially funded.			

Note. Questions of Risk of Bias assessment based on RoB2 (Sterne et al., 2019). Questions are answered for each included study to determine the risk of bias by inclusion in the analysis.

Table 24

Interpretation of Risk of Bias Assessment of Included Studies (Study 3)

Overall Risk of Bias judgement	Interpretation	Criterion
Low risk of bias	The study is comparable to a well-performed randomized trial.	The study is judged to be at **low risk of bias for all domains** for this result.
Moderate risk of bias	The study appears to provide sound evidence for a non-randomized study but cannot be considered comparable to a well-performed randomized trial.	The study is judged to be at **low or moderate risk of bias for all domains.**
Serious risk of bias	The study has one or more important problems.	The study is judged to be at **serious risk of bias** in at least one domain, but not at critical risk of bias in any domain.
Critical risk of bias	The study is too problematic to provide any useful evidence and should not be included in any synthesis.	The study is judged to be at **critical risk of bias in at least one domain.**

Note. Interpretations following RoB2 (Sterne et al., 2019).

Table 25

Overview of Risk of Bias Assessment of Included Studies (Study 3)

Study	1a	2a	2b	2c[a]	3a	3b	3c	4a	5a	6a
Adams et al. (2019)	yes	yes	yes	1	yes	yes	yes	yes	no	yes
Agostinho et al. (2015)	yes	no	n/a	n/a	yes	no	n/a	yes	no	yes
Atkinson et al. (2022)	yes	no	n/a	n/a	yes	no	n/a	yes	no	no
Bainbridge et al. (2022)	yes	yes	yes	2	yes	no	n/a	yes	yes	yes
Barenberg et al. (2021)	yes	yes	yes	18(200)	yes	no	n/a	yes	no	no
Barron (2000)	yes	yes	yes	0	yes	no	n/a	yes	no	yes
Basile (2000)	no	no	n/a	n/a	yes	no	n/a	yes	no	no
Beeson (1981)	yes	yes	yes	26	yes	no	n/a	yes	no	no
Beilstein (2021)	yes	yes	yes	0	yes	no	n/a	yes	yes	no
Bergey, Cromley, Kirchgessner & Newcombe (2015)	yes	yes	yes	11(128)	yes	yes	yes	yes	yes	yes
Bergey, Cromley & Newcombe (2015)	yes	yes	yes	8(67) / 16(115)	yes	yes	yes	yes	yes	yes
Bottge (1999)	yes	yes	yes	1	yes	yes	yes	yes	yes	yes
Boucheix & Guignard (2005)	yes	no	n/a	n/a	yes	no	n/a	yes	no	yes
Britt & Aglinskas (2002, Exp. 3)	yes	no	n/a	n/a	yes	no	n/a	yes	no	yes
Carbonneau & Marley (2015)	yes	no	n/a	n/a	yes	yes	no	yes	no	yes
Carnine et al. (1982)	yes	no	n/a	n/a	yes	yes	no	yes	yes	yes
Casey et al. (2008)	yes	no	n/a	n/a	yes	yes	yes	yes	yes	no
Chen (1999)	yes	no	n/a	n/a	yes	yes	yes	yes	no	yes
Chen (2019)	yes	yes	no	42(122)	yes	yes	yes	yes	yes	yes
Cheng (2019)	yes	yes	yes	2(90)	yes	no	n/a	yes	yes	yes
Ching & Wu (2019)	yes	yes	yes	2(140)	yes	yes	yes	yes	yes	no
Chiu & Mok (2017)	yes	yes	yes	6(129)	yes	no	n/a	yes	yes	yes
Clarke et al. (1969)	yes	yes	no	13(43)	yes	no	n/a	yes	no	no
Clements et al. (2019)	yes	yes	yes	7(152)	yes	no	n/a	yes	no	no

Table 25 (continued)

Study	1a	2a	2b	2c[a]	3a	3b	3c	4a	5a	6a
Clifton (1985)	yes	yes	yes	0	yes	yes	yes	yes	yes	no
Colón-Acosta (2019)	yes	yes	no	38(68)	yes	no	n/a	yes	yes	yes
Cramer et al. (2002)	yes	no	n/a	n/a	yes	yes	yes	no	no	no
Cromley, Bergey, et al. (2013)	yes	yes	yes	10(153)	yes	yes	yes	yes	yes	yes
Cromley, Perez, et al. (2013)	no	no	n/a	n/a	yes	yes	yes	yes	yes	yes
Day & Cordón (1993)	yes	yes	yes	0	yes	no	n/a	yes	yes	yes
DeMarinis (2011)	yes	yes	yes	0	yes	no	n/a	yes	yes	yes
Douvis (2005)	yes	no	n/a	n/a	yes	no	n/a	yes	yes	yes
Du & Zhang (2019)	yes	yes	yes	4(94)	yes	yes	yes	yes	no	no
Durnin et al. (1997)	yes	yes	yes	2(28)	yes	n/a	n/a	yes	yes	yes
Fang et al. (2016)	no	yes	yes	12(105)	no	yes	yes	yes	yes	yes
Farkas (2003)	yes	yes	yes	0	yes	yes	yes	yes	no	yes
Fennema (1972)	yes	no	n/a	n/a	yes	yes	yes	yes	no	yes
Francis (1975)	yes	no	n/a	n/a	yes	no	n/a	yes	no	no
Fuchs et al. (2003)	yes	no	n/a	n/a	yes	yes	yes	yes	yes	no
Fuchs et al. (2010)	yes	yes	yes	30(180)	yes	yes	yes	yes	yes	yes
Fuchs et al. (2016)	yes	yes	yes	18(236)	yes	yes	yes	yes	yes	yes
Fyfe (2016, Exp. 1, Exp. 3)	yes	yes	yes	1(64) / 1(50)	yes	no	n/a	yes	no	yes
Fyfe & Rittle-Johnson (2016)	yes	yes	yes	11(88)	yes	yes	yes	yes	no	no
Fyfe et al. (2015)	yes	yes	yes	10(74)	yes	yes	yes	yes	yes	no
Gellert et al. (2021)	yes	yes	yes	16(348)	yes	yes	yes	yes	yes	no
Gerjets et al. (2008)	yes	yes	yes	0	yes	no	n/a	yes	no	no
Ginns et al. (2016)	yes	yes	yes	0	yes	no	n/a	yes	no	yes
Glogger-Frey et al. (2015, Exp. 2)	yes	yes	yes	0	yes	no	n/a	yes	no	yes
Glogger-Frey et al. (2017)	yes	yes	yes	0	yes	yes	yes	yes	no	yes
Gray (2017)	yes	no	n/a	n/a	yes	no	n/a	yes	yes	no

Table 25 (continued)

Study	1a	2a	2b	2c[a]	3a	3b	3c	4a	5a	6a
Gredin & Williams (2016)	yes	yes	yes	0	yes	no	n/a	yes	yes	yes
Griffin (1995)	yes	yes	yes	5(54)	yes	yes	yes	yes	no	yes
Hoogerheide et al. (2014)	yes	yes	no	16(78)	yes	yes	yes	yes	yes	yes
Hsu et al. (2015)	yes	yes	yes	0	yes	yes	yes	yes	no	yes
Jaakkola & Veermans (2018)	yes	yes	yes	7(134)	yes	yes	yes	yes	no	no
Jaakkola & Veermans (2020)	yes	yes	yes	0	yes	yes	yes	yes	yes	yes
Johnson et al. (2014)	yes	no	n/a	n/a	yes	yes	yes	yes	no	no
Jones & Reutzel (2015)	yes	no	n/a	n/a	no	n/a	n/a	yes	yes	yes
Kaiser et al. (2018)	yes	yes	yes	35(168)	yes	yes	yes	yes	no	no
Kame'enui et al. (1986)	yes	yes	yes	3(26)	yes	no	n/a	yes	no	yes
Kaminski & Sloutsky (2020)	yes	yes	yes	2(29)	yes	no	n/a	yes	no	no
Kapur (2010)	yes	no	no	n/a	yes	yes	yes	no	yes	no
Kapur (2011)	yes	no	n/a	n/a	yes	yes	yes	yes	yes	yes
Kapur (2012)	yes	yes	yes	0	yes	yes	yes	yes	no	yes
Kapur (2014)	yes	yes	yes	0	yes	yes	yes	yes	yes	no
Kapur (2015)	no	no	n/a	n/a	yes	no	n/a	yes	no	yes
Kapur (2018)	yes	yes	yes	0	yes	yes	yes	yes	no	yes
Kapur & Kinzer (2009)	yes	no	n/a	n/a	yes	yes	yes	yes	no	no
Katzlberger (2005)	yes	yes	yes	4(49)	yes	no	n/a	yes	yes	no
Keselmann (2001)	yes	yes	yes	5(74)	yes	no	n/a	yes	yes	yes
Kirschner et al. (2009)	yes	no	n/a	n/a	yes	no	n/a	yes	no	yes
Kirschner et al. (2011)	yes	yes	yes	9(140)	yes	yes	yes	yes	yes	yes
Klahr & Nigam (2004)	yes	yes	yes	8(112)	yes	yes	yes	yes	no	no
Kneppers et al. (2007)	yes	yes	yes	0	yes	yes	no	yes	yes	yes
Kramarski & Mevarech (2003)	yes	yes	yes	0	yes	yes	yes	yes	yes	yes
Kruit et al. (2018)	yes	yes	no	302	yes	yes	yes	yes	no	yes
Kühl & Münzer (2019)	no	yes	yes	2	yes	yes	yes	yes	no	no

Table 25 (continued)

Study	1a	2a	2b	2c[a]	3a	3b	3c	4a	5a	6a
Kyun & Lee (2009)	yes	yes	yes	15(112)	yes	no	n/a	yes	no	no
Lamnina & Chase (2021)	yes	yes	yes	0	yes	no	n/a	yes	no	yes
Lawton et al. (1984)	yes	yes	yes	0	yes	no	n/a	yes	yes	no
Lehrer et al. (1999)	yes	no	n/a	n/a	yes	yes	yes	yes	no	no
Leopold et al (2013)	yes	no	n/a	n/a	yes	yes	yes	yes	no	no
Li (2013)	yes	yes	yes	6(129)	yes	yes	yes	yes	no	yes
Likourezos & Kalyuga (2017)	yes	yes	yes	0	yes	yes	yes	yes	no	yes
Lindgren et al. (2021)	yes	yes	no	17(51)	yes	yes	yes	yes	yes	no
Lonigan et al. (1999)	yes	yes	yes	15(110)	yes	yes	yes	yes	yes	yes
López et al. (1999)	yes	yes	yes	5(107)	yes	yes	yes	yes	no	yes
Magner et al. (2014)	yes	yes	yes	0	yes	yes	yes	yes	no	yes
Makransky et al. (2021)	yes	no	n/a	n/a	yes	yes	yes	yes	no	yes
Malkiewich & Chase (2019, Study 2)	yes	no	no	n/a	yes	yes	yes	yes	yes	yes
Mason et al. (2015)	yes	yes	no	11(53)	yes	yes	yes	yes	yes	no
Mason et al. (2016)	yes	yes	yes	0	yes	yes	yes	yes	no	no
Matthews & Rittle-Johnson (2009, Exp. 2)	yes	yes	yes	1(48)	yes	yes	yes	yes	no	no
McEldoon et al. (2013)	yes	yes	yes	11(80)	yes	yes	yes	no	yes	no
McHugh et al. (2021)	yes	no	n/a	n/a	yes	yes	yes	no	no	yes
McNeil et al. (2019)	yes	yes	yes	11(153)	yes	yes	yes	yes	no	no
Melo & Miranda (2015)	no	yes	yes	0	yes	yes	yes	yes	no	yes
Meneses et al. (2018)	yes	yes	yes	0	yes	yes	yes	yes	no	no
Moore & MacArthur (2012)	yes	yes	yes	9(87)	yes	yes	yes	yes	no	yes
Moreno et al. (2001)	yes	yes	yes	1(48)	yes	no	n/a	yes	no	no
Moreno et al. (2011)	yes	yes	yes	7(71)	yes	yes	yes	yes	yes	no
Mullins et al. (2011)	yes	no	n/a	n/a	yes	yes	yes	yes	yes	no

Table 25 (continued)

Study	1a	2a	2b	2c[a]	3a	3b	3c	4a	5a	6a
Muthukrishna & Borkowski (1995)	yes	yes	yes	2(106)	yes	yes	yes	yes	no	yes
Ngu et al. (2009)	yes	yes	yes	0	yes	no	n/a	yes	yes	no
Ngu et al. (2015)	no	yes	yes	0	yes	no	n/a	yes	no	yes
Nunes et al. (2009)	yes	yes	yes	0	yes	no	n/a	yes	yes	no
Obiagu et al. (2020)	no	yes	yes	0	yes	yes	yes	yes	yes	yes
Oostdam et al. (2015, Exp. 1)	yes	yes	yes	17(143)	yes	yes	yes	yes	yes	no
Paek (2012)	yes	yes	yes	11(190)	yes	yes	yes	yes	no	yes
Papaevripidou et al. (2007)	no	yes	yes	0	yes	yes	yes	yes	no	yes
Petersen et al. (2020)	yes	yes	yes	0	yes	yes	yes	yes	no	yes
Philippakos & MacArthur (2016)	yes	yes	yes	5(145)	yes	yes	yes	yes	yes	yes
Pritchard et al. (2008)	yes	yes	no	17(64)	yes	yes	yes	yes	yes	yes
Purnama & Retnowati (2021)	yes	yes	yes	0	yes	no	n/a	yes	no	yes
Rau et al. (2015, Exp. 1 & Exp. 2)	yes	yes	yes / no	20(132) / 107(259)	yes	no	n/a	yes	yes	no
Reitsma & Wesseling (1998)	yes	yes	yes	0	yes	yes	yes	yes	no	no
Ren & Gunderson (2021)	yes	yes	yes	0	yes	no	n/a	yes	no	no
Renkl et al. (2002)	yes	yes	yes	0	yes	yes	yes	yes	yes	yes
Rittle-Johnson (2006)	yes	yes	no	36(121)	yes	yes	yes	yes	no	no
Rittle-Johnson et al. (2008)	yes	yes	yes	13(67)	yes	no	n/a	yes	no	yes
Ross (1991)	yes	yes	yes	1(58)	yes	no	n/a	yes	no	yes
Safadi & Yerushalmi (2014)	yes	yes	yes	0	yes	no	n/a	yes	yes	yes
Schalk et al. (2018)	yes	yes	yes	24(213)	yes	yes	yes	yes	no	yes
Scheiter et al. (2010)	no	yes	yes	0	yes	no	n/a	yes	no	no
Scheiter et al. (2014)	no	yes	yes	0	yes	yes	yes	yes	no	no
Sidney & Alibali (2015)	yes	yes	yes	11(100)	yes	yes	yes	yes	no	yes
Smith & Smith (2012)	yes	yes	yes	2(53)	yes	no	n/a	yes	yes	no

Table 25 (continued)

Study	1a	2a	2b	2c[a]	3a	3b	3c	4a	5a	6a
Star & Rittle-Johnson (2007)	yes	yes	yes	0	yes	yes	yes	yes	yes	no
Star & Rittle-Johnson (2009)	yes	yes	yes	3(160)	yes	yes	yes	yes	yes	no
Tajika et al. (2007)	yes	yes	yes	0	yes	no	n/a	yes	no	no
Tang et al. (2019)	yes	yes	yes	2(46)	yes	yes	yes	yes	no	yes
Terwel et al. (2009)	yes	yes	yes	0	yes	yes	yes	yes	no	yes
Travlos (2010)	yes	yes	no	48(120)	yes	no	n/a	yes	yes	yes
Vadasy et al. (2015)	yes	yes	yes	38(245)/ 37(199)	yes	yes	yes	yes	yes	yes
Van Eck & Dempsey (2002)	yes	yes	no	24(112)	yes	no	n/a	no	no	yes
Van Gog et al. (2008)	yes	yes	yes	0	yes	yes	yes	yes	no	no
Van Meter et al. (2006)	yes	yes	yes	0	yes	yes	no	yes	no	yes
Wagensveld et al. (2015)	yes	yes	yes	18(119)	yes	yes	yes	yes	yes	yes
Wang et al. (2021)	yes	yes	yes	0/0	yes	yes	yes	yes	no	no
Wideman & Owston (1993)	yes	yes	yes	2/69	yes	yes	yes	yes	yes	yes
Williams & Carnine (1981)	yes	no	n/a	n/a	yes	no	n/a	no	yes	yes
Wong et al. (2002)	yes	yes	yes	0	yes	no	n/a	yes	no	no
Wouters et al. (2010)	yes	yes	yes	0	yes	no	n/a	yes	no	no
Wulf et al. (2014)	no	yes	yes	0	yes	no	n/a	yes	no	no
Yang et al. (2016)	yes	yes	yes	0	yes	no	n/a	yes	no	yes
Zhang (2019)	yes	yes	yes	0	yes	yes	yes	yes	no	no
Zhang et al. (2020)	yes	yes	yes	11(76)	yes	no	n/a	yes	no	yes

Note. Overview of results in the Risk of Bias assessment for each study. The columns refer to the questions about the Risk of Bias assessment (see Table S3). yes - The respective question is answered with a yes for this study; no - The respective question is answered with a no for this study; n/a - The question does not apply to this study.

[a] – The exact drop-out rate in numbers with the total number of participants in brackets.

Table 26

Result of Risk of Bias Assessment of Included Studies (Study 3)

Study	Baseline Confounding	Sample selection bias	Bias in measurement	Selective reporting of outcomes	Longitudinal assessment	External commercial funding	Overall rating
Adams et al. (2019)	Low risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Agostinho et al. (2015)	Low risk	Moderate risk	Moderate risk	Low risk	Moderate risk	yes	Moderate risk
Atkinson et al. (2022)	Low risk	Moderate risk	Moderate risk	Low risk	Moderate risk	no	Moderate risk
Bainbridge et al. (2022)	Low risk	Low risk	Moderate risk	Low risk	Low risk	yes	Moderate risk
Barenberg et al. (2021)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	no	Moderate risk
Barron (2000)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	yes	Moderate risk
Basile (2000)	Moderate risk	Moderate risk	Moderate risk	Low risk	Moderate risk	no	Moderate risk
Beeson (1981)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	no	Moderate risk
Beilstein (2021)	Low risk	Low risk	Moderate risk	Low risk	Low risk	no	Moderate risk
Bergey, Cromley, Kirchgessner & Newcombe (2015)	Low risk	Low risk	Low risk	Low risk	Low risk	yes	Low risk
Bergey, Cromley & Newcombe (2015)	Low risk	Low risk	Low risk	Low risk	Low risk	yes	Low risk
Bottge (1999)	Low risk	Moderate risk	Low risk	Low risk	Low risk	yes	Moderate risk
Boucheix & Guignard (2005)	Low risk	Moderate risk	Moderate risk	Low risk	Moderate risk	yes	Moderate risk
Britt & Aglinskas (2002, Exp. 3)	Low risk	Moderate risk	Moderate risk	Low risk	Moderate risk	yes	Moderate risk
Carbonneau & Marley (2015)	Low risk	Moderate risk	Moderate risk	Low risk	Moderate risk	yes	Moderate risk
Carnine et al. (1982)	Low risk	Moderate risk	Moderate risk	Low risk	Low risk	yes	Moderate risk
Casey et al. (2008)	Low risk	Moderate risk	Low risk	Low risk	Low risk	no	Moderate risk
Chen (1999)	Low risk	Moderate risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Chen (2019)	Low risk	Low risk	Low risk	Low risk	Low risk	yes	Low risk
Cheng (2019)	Low risk	Low risk	Moderate risk	Low risk	Low risk	yes	Moderate risk
Ching & Wu (2019)	Low risk	Low risk	Low risk	Low risk	Low risk	no	Low risk
Chiu & Mok (2017)	Low risk	Low risk	Moderate risk	Low risk	Low risk	yes	Moderate risk
Clarke et al. (1969)	Low risk	Moderate risk	Moderate risk	Low risk	Moderate risk	no	Moderate risk

Table 26 (continued)

Study	Baseline Confounding	Sample selection bias	Bias in measurement	Selective reporting of outcomes	Longitudinal assessment	External commercial funding	Overall rating
Clements et al. (2019)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	no	Moderate risk
Clifton (1985)	Low risk	Low risk	Low risk	Low risk	Low risk	no	Low risk
Colón-Acosta (2019)	Low risk	Moderate risk	Moderate risk	Low risk	Low risk	yes	Moderate risk
Cramer et al. (2002)	Low risk	Moderate risk	Low risk	Moderate risk	Moderate risk	no	Moderate risk
Cromley, Bergey, et al. (2013)	Low risk	Low risk	Low risk	Low risk	Low risk	yes	Low risk
Cromley, Perez, et al. (2013)	Moderate risk	Moderate risk	Low risk	Low risk	Low risk	yes	Moderate risk
Day & Cordón (1993)	Low risk	Low risk	Moderate risk	Low risk	Low risk	yes	Moderate risk
DeMarinis (2011)	Low risk	Low risk	Moderate risk	Low risk	Low risk	yes	Moderate risk
Douvis (2005)	Low risk	Moderate risk	Moderate risk	Low risk	Low risk	yes	Moderate risk
Du & Zhang (2019)	Low risk	Low risk	Low risk	Low risk	Moderate risk	no	Moderate risk
Durnin et al. (1997)	Low risk	Low risk	Moderate risk	Low risk	Low risk	yes	Moderate risk
Fang et al. (2016)	Moderate risk	Low risk	Low risk	Low risk	Low risk	yes	Moderate risk
Farkas (2003)	Low risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Fennema (1972)	Low risk	Moderate risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Francis (1975)	Low risk	Moderate risk	Moderate risk	Low risk	Moderate risk	no	Moderate risk
Fuchs et al. (2003)	Low risk	Moderate risk	Low risk	Low risk	Low risk	no	Moderate risk
Fuchs et al. (2010)	Low risk	Low risk	Low risk	Low risk	Low risk	yes	Low risk
Fuchs et al. (2016)	Low risk	Low risk	Low risk	Low risk	Low risk	yes	Low risk
Fyfe (2016, Exp. 1, Exp. 3)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	yes	Moderate risk
Fyfe & Rittle-Johnson (2016)	Low risk	Low risk	Low risk	Low risk	Moderate risk	no	Moderate risk
Fyfe et al. (2015)	Low risk	Low risk	Low risk	Low risk	Low risk	no	Low risk
Gellert et al. (2021)	Low risk	Low risk	Low risk	Low risk	Low risk	no	Low risk
Gerjets et al. (2008)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	no	Moderate risk
Ginns et al. (2016)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	yes	Moderate risk
Glogger-Frey et al. (2015, Exp. 2)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	yes	Moderate risk
Glogger-Frey et al. (2017)	Low risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk

Table 26 (continued)

Study	Baseline Confounding	Sample selection bias	Bias in measurement	Selective reporting of outcomes	Longitudinal assessment	External commercial funding	Overall rating
Gray (2017)	Low risk	Moderate risk	Moderate risk	Low risk	Low risk	no	Moderate risk
Gredin & Williams (2016)	Low risk	Low risk	Moderate risk	Low risk	Low risk	yes	Moderate risk
Griffin (1995)	Low risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Hoogerheide et al. (2014)	Low risk	Moderate risk	Low risk	Low risk	Low risk	yes	Moderate risk
Hsu et al. (2015)	Low risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Jaakkola & Veermans (2018)	Low risk	Low risk	Low risk	Low risk	Moderate risk	no	Moderate risk
Jaakkola & Veermans (2020)	Low risk	Low risk	Low risk	Low risk	Low risk	yes	Low risk
Johnson et al. (2014)	Low risk	Moderate risk	Low risk	Low risk	Moderate risk	no	Moderate risk
Jones & Reutzel (2015)	Low risk	Moderate risk	Moderate risk	Low risk	Low risk	yes	Moderate risk
Kaiser et al. (2018)	Low risk	Low risk	Low risk	Low risk	Moderate risk	no	Moderate risk
Kame'enui et al. (1986)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	yes	Moderate risk
Kaminski & Sloutsky (2020)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	no	Moderate risk
Kapur (2010)	Low risk	Moderate risk	Low risk	Moderate risk	Low risk	no	Moderate risk
Kapur (2011)	Low risk	Moderate risk	Low risk	Low risk	Low risk	yes	Moderate risk
Kapur (2012)	Low risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Kapur (2014)	Low risk	Low risk	Low risk	Low risk	Low risk	no	Low risk
Kapur (2015)	Moderate risk	Moderate risk	Moderate risk	Low risk	Moderate risk	yes	Moderate risk
Kapur (2018)	Low risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Kapur & Kinzer (2009)	Low risk	Moderate risk	Low risk	Low risk	Moderate risk	no	Moderate risk
Katzlberger (2005)	Low risk	Low risk	Moderate risk	Low risk	Low risk	no	Moderate risk
Keselmann (2001)	Low risk	Low risk	Moderate risk	Low risk	Low risk	yes	Moderate risk
Kirschner et al. (2009)	Low risk	Moderate risk	Moderate risk	Low risk	Moderate risk	yes	Moderate risk
Kirschner et al. (2011)	Low risk	Low risk	Low risk	Low risk	Low risk	yes	Low risk
Klahr & Nigam (2004)	Low risk	Low risk	Low risk	Low risk	Moderate risk	no	Moderate risk
Kneppers et al. (2007)	Low risk	Low risk	Moderate risk	Low risk	Low risk	yes	Moderate risk
Kramarski & Mevarech (2003)	Low risk	Low risk	Low risk	Low risk	Low risk	yes	Low risk

Table 26 (continued)

Study	Baseline Confounding	Sample selection bias	Bias in measurement	Selective reporting of outcomes	Longitudinal assessment	External commercial funding	Overall rating
Kruit et al. (2018)	Low risk	Moderate risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Kühl & Münzer (2019)	Moderate risk	Low risk	Low risk	Low risk	Moderate risk	no	Moderate risk
Kyun & Lee (2009)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	no	Moderate risk
Lamnina & Chase (2021)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	yes	Moderate risk
Lawton et al. (1984)	Low risk	Low risk	Moderate risk	Low risk	Low risk	no	Moderate risk
Lehrer et al. (1999)	Low risk	Moderate risk	Low risk	Low risk	Moderate risk	no	Moderate risk
Leopold et al (2013)	Low risk	Moderate risk	Low risk	Low risk	Moderate risk	no	Moderate risk
Li (2013)	Low risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Likourezos & Kalyuga (2017)	Low risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Lindgren et al. (2021)	Low risk	Moderate risk	Low risk	Low risk	Low risk	no	Moderate risk
Lonigan et al. (1999)	Low risk	Low risk	Low risk	Low risk	Low risk	yes	Low risk
López et al. (1999)	Low risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Magner et al. (2014)	Low risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Makransky et al. (2021)	Low risk	Moderate risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Malkiewich & Chase (2019, St. 2)	Low risk	Moderate risk	Low risk	Low risk	Low risk	yes	Moderate risk
Mason et al. (2015)	Low risk	Moderate risk	Low risk	Low risk	Low risk	no	Moderate risk
Mason et al. (2016)	Low risk	Low risk	Low risk	Low risk	Moderate risk	no	Moderate risk
Matthews & Rittle-Johnson (2009, Exp. 2)	Low risk	Low risk	Low risk	Low risk	Moderate risk	no	Moderate risk
McEldoon et al. (2013)	Low risk	Low risk	Low risk	Moderate risk	Low risk	no	Moderate risk
McHugh et al. (2021)	Low risk	Moderate risk	Low risk	Moderate risk	Moderate risk	yes	Moderate risk
McNeil et al. (2019)	Low risk	Low risk	Low risk	Low risk	Moderate risk	no	Moderate risk
Melo & Miranda (2015)	Moderate risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Meneses et al. (2018)	Low risk	Low risk	Low risk	Low risk	Moderate risk	no	Moderate risk
Moore & MacArthur (2012)	Low risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Moreno et al. (2001)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	no	Moderate risk

Table 26 (continued)

Study	Baseline Confounding	Sample selection bias	Bias in measurement	Selective reporting of outcomes	Longitudinal assessment	External commercial funding	Overall rating
Moreno et al. (2011)	Low risk	Low risk	Low risk	Low risk	Low risk	no	Low risk
Mullins et al. (2011)	Low risk	Moderate risk	Low risk	Low risk	Low risk	no	Moderate risk
Muthukrishna & Borkowski (1995)	Low risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Ngu et al. (2009)	Low risk	Low risk	Moderate risk	Low risk	Low risk	no	Moderate risk
Ngu et al. (2015)	Moderate risk	Low risk	Moderate risk	Low risk	Moderate risk	yes	Moderate risk
Nunes et al. (2009)	Low risk	Low risk	Moderate risk	Low risk	Low risk	no	Moderate risk
Obiagu et al. (2020)	Moderate risk	Low risk	Low risk	Low risk	Low risk	yes	Moderate risk
Oostdam et al. (2015, Exp. 1)	Low risk	Low risk	Low risk	Low risk	Low risk	no	Low risk
Paek (2012)	Low risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Papaevripidou et al. (2007)	Moderate risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Petersen et al. (2020)	Low risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Philippakos & MacArthur (2016)	Low risk	Low risk	Low risk	Low risk	Low risk	yes	Low risk
Pritchard et al. (2008)	Low risk	Moderate risk	Low risk	Low risk	Low risk	yes	Moderate risk
Purnama & Retnowati (2021)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	yes	Moderate risk
Rau et al. (2015, Exp. 1 & Exp. 2)	Low risk	Moderate risk	Moderate risk	Low risk	Low risk	no	Moderate risk
Reitsma & Wesseling (1998)	Low risk	Low risk	Low risk	Low risk	Moderate risk	no	Moderate risk
Ren & Gunderson (2021)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	no	Moderate risk
Renkl et al. (2002)	Low risk	Low risk	Low risk	Low risk	Low risk	yes	Low risk
Rittle-Johnson (2006)	Low risk	Moderate risk	Low risk	Low risk	Moderate risk	no	Moderate risk
Rittle-Johnson et al. (2008)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	yes	Moderate risk
Ross (1991)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	yes	Moderate risk
Safadi & Yerushalmi (2014)	Low risk	Low risk	Moderate risk	Low risk	Low risk	yes	Moderate risk
Schalk et al. (2018)	Low risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Scheiter et al. (2010)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	no	Moderate risk
Scheiter et al. (2014)	Moderate risk	Low risk	Low risk	Low risk	Moderate risk	no	Moderate risk

Table 26 (continued)

Study	Baseline Confounding	Sample selection bias	Bias in measurement	Selective reporting of outcomes	Longitudinal assessment	External commercial funding	Overall rating
Sidney & Alibali (2015)	Low risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Smith & Smith (2012)	Low risk	Low risk	Moderate risk	Low risk	Low risk	no	Moderate risk
Star & Rittle-Johnson (2007)	Low risk	Low risk	Low risk	Low risk	Low risk	no	Low risk
Star & Rittle-Johnson (2009)	Low risk	Low risk	Low risk	Low risk	Low risk	no	Low risk
Tajika et al. (2007)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	no	Moderate risk
Tang et al. (2019)	Low risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Terwel et al. (2009)	Low risk	Low risk	Low risk	Low risk	Moderate risk	yes	Moderate risk
Travlos (2010)	Low risk	Low risk	Moderate risk	Low risk	Low risk	yes	Moderate risk
Vadasy et al. (2015)	Low risk	Low risk	Low risk	Low risk	Low risk	yes	Low risk
Van Eck & Dempsey (2002)	Low risk	Low risk	Moderate risk	Moderate risk	Moderate risk	yes	Moderate risk
Van Gog et al. (2008)	Low risk	Low risk	Low risk	Low risk	Moderate risk	no	Moderate risk
Van Meter et al. (2006)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	yes	Moderate risk
Wagensveld et al. (2015)	Low risk	Low risk	Low risk	Low risk	Low risk	yes	Low risk
Wang et al. (2021)	Low risk	Low risk	Low risk	Low risk	Moderate risk	no	Moderate risk
Wideman & Owston (1993)	Low risk	Low risk	Low risk	Low risk	Low risk	yes	Low risk
Williams & Carnine (1981)	Low risk	Moderate risk	Moderate risk	Moderate risk	Low risk	yes	Moderate risk
Wong et al. (2002)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	no	Moderate risk
Wouters et al. (2010)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	no	Moderate risk
Wulf et al. (2014)	Moderate risk	Low risk	Moderate risk	Low risk	Moderate risk	no	Moderate risk
Yang et al. (2016)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	yes	Moderate risk
Zhang (2019)	Low risk	Low risk	Low risk	Low risk	Moderate risk	no	Moderate risk
Zhang et al. (2020)	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	yes	Moderate risk

Note. Overview of results in the Risk of Bias assessment for each study. The columns refer to the interpretation of the Risk of Bias assessment (see Table 24).

Authorship and Publication Status

Studies 1 to 3 are articles that will be submitted for publication to international, peer-reviewed scientific journals. Studies 1 and 3 were planned and conducted primarily by the author of this dissertation. In Study 2, the author of this dissertation was a co-author and contributed to the screening and coding of primary studies as well as revising the manuscript. The authors and publication status of the papers are listed below.

1. Study 1 (Chapter 5): **Dlugosch, A. C.**, Simacek, T., Simonsmeier, B. A., & Schneider, M. (manuscript in preparation). The Predictive Relations Between Conceptual and Procedural Knowledge in Mathematics: A Research Synthesis Using Meta-Analytic Structural Equation Modelling

2. Study 2 (Chapter 6): Simacek, T., **Dlugosch, A. C.**, Simonsmeier, B. A., & Schneider, M. (manuscript in preparation). How Strongly Do Motivational Constructs Mediate the Influence of Prior Knowledge on Posttest Knowledge? A Meta-Analytic Investigation of Moderated Mediation Effects

   Study 2 in this dissertation is a revised version of Study 1 in Simacek, T. (2022). *Knowledge Acquisition on the Learner Level: A Meta-Analysis*, *a Longitudinal Study and a Second-Order Meta-Analysis on Prerequisites*, *Processes*, *and Results of Learning* (https://doi.org/10.25353/ubtr-xxxx-2410-ba34) [Doctoral Dissertation, Trier University].

3. Study 3 (Chapter 7): **Dlugosch, A. C.**, Simonsmeier, B. A., & Schneider, M. (manuscript in preparation). Interventions to Foster Knowledge Transfer in School Students: A Meta-Analysis

Declaration of Authorship

I hereby declare that this dissertation was written by me independently and that no sources and aids other than those indicated were used. Content taken verbatim or in spirit from other works has been identified as such. Furthermore, the thesis has not been submitted in the same or similar form to any other university for the purpose of obtaining an academic degree. Study 2 was submitted in a previous form by Dr. Simacek as the first author of his dissertation at Trier University (Simacek, 2022).

Erklärung zur Dissertation

Hiermit erkläre ich, dass die vorliegende Dissertationsschrift von mir selbstständig angefertigt wurde und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet wurden. Wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte wurden als solche kenntlich gemacht. Zudem wurde die Arbeit in gleicher oder ähnlicher Form an keiner anderen Universität zur Erlangung eines akademischen Grades eingereicht. Studie 2 wurde in einer früheren Version von Dr. Simacek als Erstautor in seiner Dissertation an der Universität Trier eingereicht (Simacek, 2022).

Trier, ______________________ ______________________

Date/Datum, Aline C. Dlugosch