Vorsitz des Promotionsausschusses:

Prof. Dr. Stephan Winter

Berichterstatterinnen:

Prof. Dr. Daniela Czernochowski

Prof. Dr. Julia Karbach

Prof. Dr. Miriam Leuchter

Datum der wissenschaftlichen Aussprache:

03.12.2024

Vom Promotionsausschuss des Fachbereichs 8: Psychologie der Universität Koblenz-Landau zur Verleihung des akademischen Grades Doktor der Philosophie (Dr. phil.) genehmigte Dissertation

Cognitive correlates of science problem-solving in childhood

Dissertation

zur Erlangung des Doktorgrades der Philosophie (Dr. Phil.)

Fachbereich Psychologie

Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau (RPTU)

vorgelegt von

Jonas Schäfer, M. Sc.

Landau in der Pfalz, 2024

Table of Contents

Abstract

Problem-solving is part of many activities in every-day life and of educational tasks. Problem-solving abilities have been shown to benefit personal and professional success and, thus, are considered an essential learning outcome in early and middle childhood, particularly in the science domain. Previous research has indicated that problem-solving is based on a complex interplay of domain-specific knowledge, such as science concept knowledge, and domain-general cognitive abilities, such as reasoning. However, there is disagreement on whether domain-specific knowledge or domain-general cognitive abilities are the dominant predictor of children's problem-solving performance. Furthermore, evidence on the impact of executive functions on children's problem-solving performance is rare and inconclusive, which may be attributed to an insufficient number of appropriate assessment instruments. Therefore, the two major goals of the present dissertation were (1) to develop and validate new digital problem-solving tasks for children and (2) to investigate the cognitive correlates of science problem-solving in elementary school children. In Study 1, science problem-solving tasks with gears and building blocks were developed for tablet-based application and validated by demonstrating measurement invariance and convergent validity in a sample of 215 six- to eight-year-old children. Based on the same sample, Study 2 showed that domain-specific knowledge was significantly less predictive for science problem-solving performance than the domain-general cognitive abilities language and reasoning. Based on another sample of 476 six- to eight-year-olds, Study 3 demonstrated that the core executive functions working memory and cognitive flexibility were positively related to science problem-solving performance, but inhibition was not. These findings provide important new insights into the cognitive correlates of children's problem-solving that are discussed with regard to problem structures, cognitive mechanisms, and future research perspectives.

German Summary (Zusammenfassung)

Problemlösen ist Teil vieler Alltagsaktivitäten und schulischer Aufgaben. Es hat sich gezeigt, dass Problemlösungsfähigkeiten zu persönlichem und beruflichem Erfolg beitragen und daher, gerade im Bereich der Naturwissenschaften, als wesentliches Lernziel in der frühen und mittleren Kindheit angesehen werden. Bisherige Forschung hat gezeigt, dass das Problemlösen auf einem komplexen Zusammenspiel von bereichsspezifischem Wissen, wie dem Wissen über naturwissenschaftliche Konzepte, und bereichsübergreifenden kognitiven Fähigkeiten, wie dem logischen Denken, beruht. Es besteht jedoch Uneinigkeit darüber, ob bereichsspezifisches Wissen oder bereichsübergreifende kognitive Fähigkeiten der wichtigere Prädiktor für die Problemlösungsleistung von Kindern ist. Darüber hinaus gibt es nur wenig und widersprüchliche Evidenz für den Einfluss exekutiver Funktionen auf die Problemlösungsleistung von Kindern, was auf eine unzureichende Anzahl geeigneter Erhebungsinstrumente zurückzuführen sein könnte. Daher waren die beiden Hauptziele der vorliegenden Dissertation (1) die Entwicklung und Validierung neuer digitaler Problemlösungsaufgaben für Kinder und (2) die Untersuchung der kognitiven Korrelate des naturwissenschaftlichen Problemlösens bei Grundschulkindern. In Studie 1 wurden naturwissenschaftliche Problemlösungsaufgaben mit Zahnrädern und Bausteinen für eine tabletbasierte Anwendung entwickelt und durch den Nachweis von Messinvarianz und konvergenter Validität in einer Stichprobe von 215 sechs- bis achtjährigen Kindern validiert. Auf Grundlage derselben Stichprobe zeigte Studie 2, dass bereichsspezifisches Wissen einen signifikant geringeren Einfluss auf die Leistung im naturwissenschaftlichen Problemlösen hatte als die bereichsübergreifenden kognitiven Fähigkeiten Sprache und logisches Denken. Auf Grundlage einer weiteren Stichprobe von 476 Sechs- bis Achtjährigen zeigte Studie 3, dass die exekutiven Kernfunktionen Arbeitsgedächtnis und kognitive Flexibilität positiv mit der Leistung im naturwissenschaftlichen Problemlösen zusammenhingen, die Inhibition

hingegen nicht. Diese Ergebnisse liefern wichtige neue Erkenntnisse über kognitive Fähigkeiten, die kindlichem Problemlösen zugrunde liegen, welche im Hinblick auf Problemstrukturen, kognitive Mechanismen und zukünftige Forschungsperspektiven diskutiert werden.

1 Theoretical and Methodological Background

The cognitive correlates of children's problem-solving are subject of research in educational and developmental psychology (Keen, 2011; Mayer & Wittrock, 2006). Problem-solving abilities represent the skill-set involved in performing challenging tasks by overcoming obstacles (Duncker, 1945). In childhood, problem-solving abilities are associated with success in various areas of life, on a social and academic level (Lile Diamond, 2018). Moreover, problem-solving abilities are an essential educational outcome that is closely related to students' science, technology, engineering, and mathematics (*STEM*) learning (OECD, 2014).

Previous research suggests that a broad range of cognitive abilities is involved in problem-solving (Bartley et al., 2018; Funke et al., 2018; Wang & Chiew, 2010). However, to date, only a limited number of quantitative studies on the cognitive underpinnings of children's problem-solving abilities in the science domain (i.e., *science problem-solving*) has been conducted. In particular, the relationships between children's science problem-solving performance and their domain-specific knowledge, executive functions, and intelligence have not exhaustively been investigated. Moreover, the methodology in problem-solving research is partly inconsistent between studies (cf. Grubbs et al., 2018; Jonassen, 2013). This may be the consequence of a lack of instruments for the assessment of problem-solving abilities with an educational and theoretical foundation (Greiff & Fischer, 2013).

1.1 Problem-Solving

The term problem-solving describes the process of changing an initial state to a goal state when no obvious way to achieving the goal is known to the problem solver (Duncker, 1945; Mayer, 1992). Since there are inconsistent definitions of problem-solving (cf. Schoenfeld, 2016), it is not entirely clear when a task can be considered a problem and which behavioral and cognitive processes are considered problem-solving (Dörner & Funke, 2017; Schmidt, 2011). A widely used criterion to define a problem is the absence of a routine behavior that can be employed to achieve the goal state (Funke et al., 2018). However, given that problem-solving may necessitate the implementation of automated procedures, particularly when executing strategies, it is questionable whether routines can be excluded as components of problem-solving (Mayer, 1998). For example, if a problem requires to solve a memory game, the problem solver will perform motor actions to reveal cards, which are highly automated and routine. Consequently, a lack of routine knowledge is not a sufficient criterion for a task to be considered a problem. A problem rather is a task where the desired goal state cannot be attained exclusively by the use of routines (Schoenfeld, 2016). According to Funke (2003), the gap between the initial state and the goal state, which cannot be closed through routine behavior, marks the space in which problem-solving cognition is involved. Thus, a task may not be classified as a problem solely on the basis of external factors, but it also depends on intrapersonal factors, such as the problem solver's prior knowledge and cognitive abilities (Hambrick et al., 2020; Mehadi Rahman, 2019; Wu & Molnár, 2022).

Problem-solving is considered a process that includes a number of phases (Funke, 2012; Mayer, 1992; Pólya, 1957). Previous studies have provided a range of disparate descriptions and numbers of problem-solving phases (cf. Priemer et al., 2019), which may be subsumed under the following four main phases (oriented to OECD, 2013; Pólya, 1957; Zelazo et al., 1997): a) Understanding and representing the problem, b) planning and developing strategies, c) executing strategies, and d) evaluating and monitoring. Phases a) and b) require mental processes and the exploration of the problem space (Eichmann et al., 2019; Rowe et al., 2001). Phases c) and d) involve actively dealing with the problem, e.g., by observing, testing, and optimizing intermediate solutions (Gold et al., 2021). The problem-solving phases are regarded as distinct but interdependent, whereby progress achieved in one specific phase may benefit progress in another phase (Eichmann et al., 2019). The problem-solving process may be iterative, i.e., the phases may be repeated several times (Hollenstein et al., 2022). Switching between phases (i.e., transitioning) can be cognitively demanding but may have a positive effect on problem-solving performance (Molnár & Greiff, 2023).

1.1.1 The structure of a problem

Well- and ill-defined problems. An essential feature of problem structures is the clarity and transparency of their goal states, which can be classified as either well-defined or ill-defined (synonymously called well- or ill-structured; Allaire & Marsiske, 2002). A well-defined problem provides a unique goal state and, in consequence, an unambiguously correct solution that can be achieved by a limited number of solution pathways (Schraw et al., 1995). A well-defined problem may, for example, require the construction of a box with a hole, through which bees may enter but frogs may not (cf. Strimel et al., 2018). An ill-defined problem may provide different conflicting goal states that may initially be non-transparent to the problem solver, but have to be discovered (Funke, 2003). The number of possible solutions in ill-defined problems may be large or even unlimited and there might be no way of solving it optimally (Schraw et al., 1995). Instead, an optimal solution of an ill-defined problem might be ambiguous and depend on the problem solver's perspective (Lynch et al., 2009; Sarathy, 2018). An ill-defined problem might, for instance, require to solve societal or political issues (Jonassen, 1997).

The criteria that a problem needs to meet to be classified as well- or ill-defined are controversial and the degree to which problems can be ill-defined is potentially unlimited (Lynch et al., 2009). Therefore, the differentiation between well- and ill-defined problems should not be seen as a dichotomous categorization, but as a continuum (Jonassen, 2013).

Nevertheless, the clarity and transparency of a problem goal state has important implications for assessment methods of problem-solving that will be discussed in section 1.1.3.

Complex problems. A *complex problem* describes a task with a high degree of interdependencies between task parameters or task states, i.e., initial state, current state, and goal state (Dörner & Funke, 2017; Frensch & Funke, 1995). For example, interdependencies may relate to properties of task-relevant objects (e.g., an object only moves when it is pushed by another object) or rule-based principles (e.g., changing the state of one parameter also affects the states of other parameters; Greiff et al., 2016). Greiff et al. (2016) assumed that complex problems include two dimensions: Rule knowledge, which refers to the acquisition of knowledge about the interdependencies underlying a problem, and rule application, which refers to the application of this knowledge to achieve solutions. According to Funke (2010), the amount of higher-order cognition that is required to solve a problem raises as a function of the problem complexity. This is consistent with the findings that reasoning contributes to the ability to cognitively segment complex problems into component parts in elementary school children (O'Brien et al., 2023) and adults (Duncan et al., 2017). Consequently, complex problems have a theoretical overlap with ill-defined problems, because they both require to gather new information to solve them. However, an important difference is that ill-defined problems entail uncertainties and a high non-transparency, while complex problems may provide a clear goal state, but incorporate complicated interdependencies (Jonassen, 2013).

Problem paradigms. A meta-analysis of problem-solving experiments identified three problem-solving paradigms, namely the mathematical, verbal, and visuospatial paradigms (Bartley et al., 2018). These paradigms encompassed a number of distinct task types, including *insight problem* tasks, *deductive*, *inductive*, and *visuospatial relational reasoning* tasks (Bartley et al., 2018). An insight problem describes a problem in which the recognition of a mechanism, principle, or rule during the problem-solving process can facilitate immediate progress (Gaschler, 2020). This insight prompts a shift of the problem solver's attentional focus to aspects of the problem that are relevant for developing an effective strategy (Korovkin et al., 2018). According to Bartley et al. (2018), reasoning tasks are defined as follows: Deductive reasoning tasks involve drawing specific conclusions from general rules, inductive reasoning tasks involve inferring general rules from specific instances and experiences, and visuospatial relational reasoning tasks involve understanding spatial patterns and interdependencies by identifying or verifying rules.

1.1.2 Children's problem-solving in the STEM domains

Problem-solving abilities develop during early and middle childhood because of improvements in planning (Injoque-Ricle et al., 2014) and in the goal-directed use of task-relevant objects (Keen, 2011). However, when planning skills are not yet fully developed, children may use trial-and-error strategies that may still lead to satisfactory solutions (Tönnsen, 2021). Educational research has identified problem-solving abilities as an essential component in STEM learning (Astuti et al., 2021; English, 2023). In the context of STEM learning, problem-solving abilities are considered as a competency that goes beyond factual knowledge (Greiff & Fischer, 2013) and can be assessed using task-relevant objects that motivate children to engage with, e.g., building blocks (Weber & Leuchter, 2020). Identifying the processes underlying problem-solving abilities in children in the STEM domains requires considering the acquisition of relevant science concepts.

The acquisition of science concepts during childhood. Science concept knowledge is acquired as a result of experiences that children make in their natural environment (Fragkiadaki et al., 2023). Although young children's initial science concepts are mostly inaccurate, they still attempt to find regularities in science phenomena they perceive (Wilkening & Cacchione, 2014). Consequently, young children mentally represent naïve pre-concepts that might be based on *phenomenological primitives* (i.e., small pieces of experience-based knowledge that are not yet coherently linked; diSessa, 1988). According to Leuchter and Hardy (2021), the acquisition of accurate science concept knowledge (i.e., a *conceptual change*) is based on children's scientific thinking and the iterative comparison of theory and evidence. Since science concepts are typically based on rule-based structures (e.g., when the center-of-mass of an object is supported it is stable and vice versa; Weber & Leuchter, 2020), deductive, inductive, and abductive reasoning are essential for children to acquire science concept knowledge (Leuchter & Hardy, 2021). Children's reasoning is also a cognitive prerequisite for other important science skills, such as scientific argumentation and designing unconfounded experiments (Leuchter & Hardy, 2021).

In the realm of problem-solving in the STEM domains, science concept knowledge has been operationalized as domain-specific knowledge that might affect children's problem-solving performance and process (Reuter & Leuchter, 2021; Strimel et al., 2018). Important concepts that children typically acquire during pre- and elementary school age relate to the turning dynamics of gears (Lehrer & Schauble, 1998; Reuter & Leuchter, 2021) and the stability of asymmetric building block constructions (Krist, 2010; Weber & Leuchter, 2020; Weber et al., 2020). The turning dynamics of gears (i.e., turning direction and turning speed) represent rule-based aspects of the transmission of motion (Lehrer & Schauble, 1998). The turning direction of adjacently interconnected gears is opposite, which implies that two gears turn in the same direction when they are connected by an odd number of intermediate gears (Reuter & Leuchter, 2021). The turning speed of gears depends on their size, such that larger gears turn slower around their axis than smaller gears (Reuter & Leuchter, 2021). Lehrer and Schauble (1998) found that approximately 70 % of second-graders and approximately 90% of fifth-graders made accurate judgements regarding the turning direction of interconnected gears. The percentage of accurate judgements regarding gear turning speed was approximately 50% in both second- and fifth-graders (Lehrer & Schauble, 1998). Reuter and Leuchter (2021) found similar percentages of accurate judgements that increased as a function of age, from five-year-olds (turning direction: 39%; turning speed: 33%) to ten-year-olds (turning direction: 80%; turning speed: 43%). The most prevalent naïve pre-concepts were that interconnected gears always turn in the same direction and that larger gears turn faster (Reuter & Leuchter, 2021). The stability of asymmetric constructions represents a rule-based aspect of statics principles (Krist, 2010). Precisely, the stability of asymmetric building block constructions depends on whether the center-of-mass of the building blocks is supported (Weber & Leuchter, 2020). Krist (2010) found that three- to six-year-olds' judgements regarding the stability of building block constructions were significantly more accurate when the constructions were symmetrical than when they were asymmetrical. Weber and Leuchter (2020) found that less than 50% of five- and six-year-old children accurately judged the stability of asymmetrical building block constructions in both stable and instable items. These results suggest that most preschool children hold the naïve pre-concept that stability depends on the geometrical center (Weber & Leuchter, 2020).

The relationship of science concept knowledge and problem-solving is bidirectional because problem-solving may provide insights to the problem solver that his or her prior knowledge does not match what is being experienced (i.e., cognitive conflict) and needs to be assimilated (Yeo & Tan, 2014). Consequently, cognitive conflicts, which are caused by a problematic task a child faces, may initiate a shift from naïve pre-concepts to more elaborate post-concepts (i.e., *conceptual change*; Demetriou et al., 2011; Piaget, 1977). The research paradigm investigating problem-solving as a learning vehicle is called problem-based learning or challenge-based learning (Yeo & Tan, 2014).

1.1.3 Assessment methods

A variety of tasks has been employed for the assessment of different aspects of problem-solving abilities (e.g., Greiff et al., 2015a; Wicaksono & Korom, 2022), with a key distinction between measures of problem-solving performance and measures of the problem-solving process (Jonassen, 2013).

Types of measures. Problem-solving performance is typically assessed in terms of the solution quality, which refers to how accurate a solution meets given goal criteria (Jonassen, 2013). Thus, the solution state generated by a problem solver (i.e., a participant) in a problem-solving task is compared with an optimal solution (i.e., a reference solution) according to the relevant goal criteria (Allaire & Marsiske, 2002). For example, it would be evaluated whether a hole in a box was constructed sufficiently large to let bees enter but small enough to exclude frogs (Strimel et al., 2018).

Process measures, on the other hand, have mainly been implemented by using video-based observations and protocols of the participant's behavior throughout the problem-solving process (e.g., Ramey & Uttal, 2017; Strimel et al., 2018). For instance, Strimel et al. (2018) collected *concurrent think-aloud protocols* of preschoolers and fourth-graders who solved engineering design problems and were asked to simultaneously verbalize their thoughts. The authors evaluated these protocols according to a coding scheme consisting of 17 mental processes for technological problem-solving (Halfin, 1973; Strimel et al., 2018; Wicklein & Rojewski, 1999). Although think-aloud protocols suggest an immediate path to the identification of mental processes, the validity of such self-reports of young children has been doubted (Conijn et al., 2020). Moreover, concurrent think-aloud protocoling has been shown to diminish students' performance in problem detection tasks (Van den Haak et al., 2003). Another method to analyze the problem-solving process is *observation protocoling*, where the participant's behavior is documented video-based or by an independent observer (Hidayati & Wagiran, 2020). Observation protocols reflect self-determined behavior of the participant, without eliciting verbal behavior, as in think-aloud settings (Smith et al., 2013; Van den Haak et al., 2003). Observation protocols can yield quantitative data, such as number, timing, and type of the participant's actions (e.g., manipulating task-relevant objects), that are an effective indicator of the problem-solving process (Molnár & Greiff, 2023). However, they can also be used to categorize and interpret behavior qualitatively, e.g., by coding the intentions, strategies, or motivation of the participant (Grubbs et al., 2018). Qualitative analyses of the problem-solving process have been criticized for being hardly comparable across studies because many different kinds of coding schemes have been applied across studies (Grubbs et al., 2018).

Investigations of well-defined problem-solving provide the advantage that solution quality can objectively indicate task performance (Allaire & Marsiske, 2002). Therefore, the term problem-solving performance is used to refer to solution quality in the following chapters. Investigations of ill-defined problem-solving typically focus on the problem-solving process, since ill-defined problems do not provide objectively correct reference solutions (e.g., Schunn et al., 2005).

Task modalities. Traditionally, problem-solving has usually been assessed in analog settings (e.g., Chan, 1989; Shallice, 1982). For instance, a classical well-defined problem-solving task, the *Tower of London*, was conducted using three wooden or plastic balls that needed to be arranged on a board with three pegs (Shallice, 1982). The goal of the Tower of London task is to match the pattern of the balls to a pictured arrangement, where no more than two balls may be placed on the middle peg, and no more than one ball may be placed on the smallest peg (Keith Berg & Byrd, 2002). Analog problem-solving assessments provide the advantage that participants may be tested in a natural environment (Shechter et al., 2021). Nowadays, the majority of problem-solving studies uses digital assessment instruments (e.g., Gao et al., 2022; Wu & Molnár, 2022). A prominent paradigm assessing problem-solving abilities digitally are *Microworlds*, such as the *MicroDYN*, that simulate real-world problems (Bühner et al., 2008; Greiff et al., 2016; Molnár & Csapó, 2018). In the MicroDYN approach, participants are presented with fictive scenarios on a computer screen that, for example, require them to solve ill-defined problems, such as helping a cat to recover from sickness (Wu & Molnár, 2022). Digital assessment instruments provide several advantages over analog instruments by facilitating standardized experimentation and enabling more accurate, efficient, and reliable data coding (Germine et al., 2019; Greiff et al., 2015b).

Tracking data. A central benefit of digital assessment is the output of tracking data, which yields an efficient method for analyzing problem-solving abilities (Csapó et al., 2012; Greiff et al., 2015b; Molnár & Csapó, 2018; Wu & Molnár, 2022). Tracking data generated during problem-solving on digital devices provides rich opportunities to study both the problem-solving performance, by automated solution matching (e.g., Molnár & Csapó, 2018), and the problem-solving process, e.g., by analyzing the frequency and rhythm of executed actions (e.g., Wu & Molnár, 2022). A popular large-scale sample of problem-solving tracking data was collected within the *Programme for International Student Assessment* (*PISA*) of the *Organisation for Economic Co-operation and Development* (*OECD*; OECD, 2014). Several researchers analyzed the PISA tracking data to identify behavioral patterns in complex problem-solving of 15-year-old students (e.g., Eichmann et al., 2019; Greiff et al., 2015b). Since tracking data, such as log-files, are protocols of actions executed in a computer program, they are typically output in computer code format that may be difficult to interpret (Csapó et al., 2012). Making tracking data analyzable for research purposes may require pre-processing by complex algorithms, particularly when pre-processing steps are not carried out automatically by the test software (Greiff et al., 2015b). Furthermore, data interpretation may be challenging because many actions cannot be attributed unambiguously to the intention of a participant (Whitelock-Wainwright et al., 2020). Nevertheless, tracking data provides a measure that surpasses human coding in terms of accuracy, objectivity, and efficiency (Csapó et al., 2012).

Despite the substantial advantages of digital assessment, it is an ongoing debate whether the modality (analog or digital) has an impact on the cognitive demands a task poses (Bignardi et al., 2021; Germine et al., 2019; Guilbert et al., 2019; Mayer, 2005). Therefore, newly developed digital assessment instruments require to be validated against their analog counterpart in a sample of the target group (Björngrim et al., 2019; Johann & Karbach, 2018; Vermeent et al., 2022).

1.2 Higher-Order Cognition

1.2.1 Intelligence

Intelligence is considered as the mental ability of humans to adapt to their environment and to learn from experience (Sternberg, 2012). Psychological research has identified a number of theoretical intelligence models (Frischkorn et al., 2022), of which the model developed by Cattell (1943) is widely established. This model includes two complementary components of general intelligence: Fluid intelligence, which describes higher-order cognitive functions for information processing, and crystallized intelligence, which describes cognitive resources that result from learning and experience (Brown, 2016; Cattell, 1943). Fluid intelligence enables humans to process new information and to handle novel situations in a goal-directed manner (Heaton et al., 2014). Crystallized intelligence represents acquired knowledge and trained skills (Harada et al., 2013). Thus, fluid intelligence is a cognitive prerequisite for learning, whereas crystallized intelligence is a product of learning (Dixon et al., 2013). However, both components of intelligence may interact in learning processes (Buades-Sitjar et al., 2022; Weber et al., 2020), for example when new information is integrated into existing knowledge structures. Fluid intelligence is considered an inherited trait that develops from birth to early adulthood and begins to decline at approximately 30 years of age, while crystallized intelligence may improve throughout the lifespan and remain stable in old age (Salthouse, 2012). Essential indicators of fluid intelligence are reasoning as well as abstract and logical thinking (Chen et al., 2019; Liu et al., 2024). Essential indicators of crystallized intelligence are acquired knowledge, facts, and certain language aspects, such as vocabulary (Buades-Sitjar & Duñabeitia, 2022). A number of well-established test batteries is available to assess intelligence (Tager-Flusberg & Plesa-Skwerer, 2009), which typically include subscales for the measurement of reasoning and language (e.g., the *Wechsler Preschool and Primary Scale of Intelligence*; Wechsler, 2012).

1.2.2 Executive functions

Executive functions are a set of cognitive control functions that enable goal-directed action and self-regulated behavior (Nigg, 2017). Therefore, they are essential for children's cognitive and social development (Diamond, 2013). Executive functions essentially improve during early childhood but keep developing into adolescence (Wiebe & Karbach, 2017). Previous studies identified interventions, such as cognitive trainings, that might improve some aspects of executive functions (Diamond & Lee, 2011; Titz & Karbach, 2014). The core executive functions are inhibition, working memory, and cognitive flexibility (Diamond, 2013; Miyake et al., 2000).

Inhibition. Inhibition refers to the cognitive mechanisms of action control that suppress undesired stimulus responses (Johnstone et al., 2007). Thus, inhibition serves to maintain limited cognitive resources (Bari & Robbins, 2013). Psychological research has distinguished two main kinds of inhibition, namely resistance to distractor interference and inhibition of prepotent responses (Rey-Mermet et al., 2018). Resistance to distractor interference describes the suppression of perceived information from a useless stimulus, while inhibition of prepotent responses describes the suppression of a motor response that is dominantly triggered by a stimulus (Rey-Mermet et al., 2018). Inhibition improves essentially during the preschool years and changes only marginally during late childhood and adolescence

(Best & Miller, 2010). Inhibition may have a positive effect on academic success because it allows children to suppress stimuli that are disruptive for learning (Zamora et al., 2020). Moreover, inhibition has been shown to contribute to self-regulation, decision-making, and creativity in children (Cassotti et al., 2016; Hofmann et al., 2012; Nigg, 2017). In tasks involving few or no task-irrelevant stimuli, inhibition of prepotent responses is a more reliable correlate than resistance to distractor interference (Lee et al., 2009). A widely established task measuring inhibition of prepotent responses in children is the *Go/No-go* task, in which participants are continuously presented with stimuli signaling whether they should rapidly respond or not (Johnstone et al., 2007). The Go/No-go-ratio lays around 3:1 in order to establish the Go-response as prepotent and evoke inhibition in No-go-trials (St. John et al., 2019). This task type offers the opportunity of a child-friendly presentation, e.g., with fairytale-like stimuli and cover stories (Johann & Karbach, 2018; St. John et al., 2019).

Working memory. Working memory describes a set of cognitive mechanisms that enables the maintenance and manipulation of perceived information within the scope of seconds to a few minutes after stimulus presentation (D'Esposito, 2008). Thus, working memory is involved in most mentally demanding tasks (Ellis et al., 2020). Among a large number of theoretical models of working memory (Adams et al., 2018), the three-component model of Baddeley and Hitch (1994) is widely accepted. It is considered to comprise the two storage systems, visuospatial sketchpad and phonological loop, as well as the central executive that manages the information transfer between modalities and the attentional control (Baddeley & Hitch, 1994). Working memory capacity is typically indicated by the *working memory span*, defined as the number of information units that a person is capable to maintain simultaneously (Cowan, 2010). The working memory span increases throughout childhood from approximately two units in four-year-olds to approximately four to five units in 14-year-olds and adults (Cowan, 2010; Gathercole et al., 2004). However, these spans may vary according to individual cognitive abilities and stimulus type (verbal, visual, auditory), and can be extended by chunking and rehearsal processes (Hurlstone et al., 2014; Thalmann et al., 2019). Working memory is involved in the formation of long-term memories (Cowan, 2014) and is, thus, closely associated with learning progress and academic achievement in childhood (Forsberg et al., 2021; Holmes & Adams, 2006; Johann & Karbach, 2020). In tasks requiring the consideration of spatial interdependencies, visuospatial working memory is a more reliable correlate than verbal working memory (Hodgkiss et al., 2018). One well-established task to assess the visuospatial working memory span is the *Corsi Blocks* task (Corsi, 1972; Gathercole et al., 2004). In this task, participants see a number of blocks, of which an increasing number is sequentially highlighted. After presentation, participants should recall the highlighted blocks in the order they were highlighted (Hurlstone & Hitch, 2018). To enhance working memory demands, studies have employed backward task versions that require to recall the blocks in reverse order of presentation (e.g., Alloway et al., 2006). This task type has been established in clinical and research settings (Farrell Pagulayan et al., 2006; Stoffers et al., 2003) and provides the opportunity of a child-friendly presentation (Johann & Karbach, 2018).

Cognitive flexibility. Cognitive flexibility (synonymously called flexibility, shifting, or task switching) refers to the ability to switch attention between different perspectives and task demands (Miyake et al., 2000; Titz & Karbach, 2014). Furthermore, cognitive flexibility enables to switch between mental concepts, such as strategies or rules, throughout a task (Diamond, 2013). Cognitive flexibility develops during early childhood (Deák & Wiseheart, 2015) but continues to improve during late childhood and adolescence (Buttelmann & Karbach, 2017). In children, cognitive flexibility is positively associated with math abilities (Yeniad et al., 2013), reading comprehension (Johann et al., 2020), and academic achievement (Titz & Karbach, 2014, for a review). One well-established way to measure cognitive flexibility in children are *Card Sorting* tasks, in which participants match visual features of presented cards according to repeatedly changing categories (Jacques & Zelazo, 2001; Ozonoff et al., 2005; Somsen, 2007; Zelazo, 2006). This task type also provides opportunities to present it in a child-friendly format by the use of appropriate card symbols (Marcus et al., 2020).

1.3 Associations of Problem-Solving and Other Cognitive Abilities

Previous research has identified associations of problem-solving abilities and higher-order cognitive abilities, such as language (e.g., Baldo et al., 2005; Gunzenhauser et al., 2019), reasoning (e.g., Greiff et al., 2016; Leighton & Sternberg, 2003), and executive functions (e.g., Viterbori et al., 2017; Zook et al., 2004).

Some aspects of language, as indicators of crystallized intelligence, have been shown to be involved in problem-solving (e.g., Chan & Kwan, 2021). In particular, self-directed speech had positive effects on performance on the Tower of London task (Gunzenhauser et al., 2019) and on planning performance (Enke et al., 2022) in elementary school children. Inner speech (i.e., the subjective experience of language without overt articulation; Alderson-Day & Fernyhough, 2015) is an instance of self-directed speech that can reduce cognitive load during problem-solving (Kompa & Mueller, 2022). In a study of Rohrkemper (1986), elementary school children reported that they used inner speech as a strategy in problem-solving tasks. Furthermore, problem-solving performance in adults declined when their inner speech was disrupted, providing further evidence for the benefits of language-based strategies (Baldo et al., 2005; Wallace et al., 2017). Previous studies have also demonstrated that the domain-specific vocabulary (e.g., mathematical terminology in mathematical word problem-solving) improves problem-solving performance in third-graders (Chan & Kwan, 2021) and older adults (Chen et al., 2017).

Reasoning, as an indicator of fluid intelligence, enables logical, scientific, and evidence-based thinking processes that may have a positive impact on problem-solving in childhood and adolescence (Kim & Pegg, 2019; Leuchter & Hardy, 2021; Mehadi Rahman, 2019; Tan et al., 2023). Especially for complex problems, reasoning has been shown to be a strong predictor of problem-solving performance (e.g., Greiff et al., 2016; Wüstenberg et al., 2012), although there is some counterevidence (Kretzschmar & Nebe, 2021). Mayer et al. (2014) have demonstrated an association of reasoning and problem-solving in elementary school children, in the way that problem-solving abilities predicted individual differences in scientific reasoning. Leighton and Sternberg (2003) concluded that reasoning and problem-solving provide many commonalities but are separable.

Some researchers have argued that intelligence and problem-solving are identical constructs (e.g., Hambrick et al., 2020), while others considered them as entirely separate (e.g., Rigas & Brehmer, 1999; see Stadler et al., 2015, for a review). There is a notable distinction between intelligence, which is a purely theoretical construct, and problem-solving, which is a behavior that might be observable (De Boeck et al., 2023; Smith et al., 2013). Nevertheless, when considering problem-solving abilities in a theoretical context, there is a broad consensus that they are closely associated with intelligence (Kretzschmar et al., 2016; Kröner et al., 2005; Stadler et al., 2015; Wüstenberg et al., 2012). This is underlined by medium to high correlations between both constructs, that have, however, been interpreted differently: Kretzschmar et al. (2016) suggested that intelligence does not explain the full variance in problem-solving abilities, demonstrating discriminant validity of problem-solving abilities. Other researchers, who found particularly strong correlations between measures of intelligence and problem-solving abilities, concluded that problem-solving abilities are an indicator of intelligence (Kröner et al., 2005; Wüstenberg et al., 2012).

The relationship between intelligence and problem-solving abilities is a topic that is closely related to the debate on the nature of domain-general problem-solving abilities (Kretzschmar & Nebe, 2021; Tricot & Sweller, 2014). Cognitive psychology has distinguished between domain-specific and domain-general cognitive resources, assuming that some cognitive resources are shared between different demands and content domains, while others are dedicated to specific demands and content domains (Demetriou et al., 2011). For instance, in studies on elementary school children's mathematical performance, working memory has been regarded as a domain-general resource that is involved in solving problems with heightened complexity (Avcil & Artemenko, 2023; Soltanlou et al., 2017). In contrast, domain-specific resources, such as counting knowledge (Träff et al., 2023), were more strongly involved in less complex problems, such as one-digit multiplication (Soltanlou et al., 2017). This distinction was evidenced by higher activation in prefrontal and lower activation in parietal brain regions during tasks that mainly required domain-general resources compared to tasks that mainly required domain-specific resources (Soltanlou et al., 2017).

Consequently, domain-specific knowledge refers to knowledge of particular aspects of the problem's content domain, such as science concepts, while higher-order cognitive abilities, such as intelligence and executive functions, are considered as domain-general (Miyake & Friedman, 2012). Some previous studies have assumed that domain-specific knowledge, also referred to as expert knowledge (*expert-novice paradigm*), is the key factor in problem-solving performance, particularly when dealing with well-defined problems (Chi et al., 1981; Perkins & Salomon, 1989). Tricot and Sweller (2014) even argued that almost any problem can be solved based on domain-specific knowledge, while domain-general abilities are rarely involved in problem-solving. However, several studies have demonstrated that domain-general abilities, particularly reasoning and working memory, were closely associated with students' problem-solving performance (Greiff et al., 2016; O'Brien et al., 2023; Swanson &

Beebe-Frankenberger, 2004). Domain-specific knowledge and domain-general abilities may interact, e.g., such that domain-general abilities enable the application of domain-specific knowledge (Perkins & Salomon, 1989; Roberts, 2007). To date, there is a lack of studies comparing the impact of domain-specific knowledge and domain-general abilities on children's problem-solving performance.

Executive functions have been shown to increase students' performance in mathematical word problem-solving (Swanson & Beebe-Frankenberger, 2004) and on the Tower of London task (Zook et al., 2004). Particularly working memory was identified as a main source of variance in students' complex (Bühner et al., 2008; Kretzschmar & Nebe, 2021) and mathematical (Viterbori et al., 2017) problem-solving performance. Inhibition and cognitive flexibility have rarely been investigated in association with problem-solving, but they did not contribute to mathematical problem-solving performance in the study of Viterbori et al. (2017). Findings of the contribution of executive functions on problem-solving performance in the science domain are sparse.

1.4 Development of the Tasks

Within the framework of this dissertation, a number of tasks assessing science problem-solving abilities, domain-specific knowledge, and executive functions were developed for tablet-based use and subsequently employed in the dissertation studies (see Table 1). The target age group for these tasks were pre- and elementary school children. In previous studies, the tasks assessing science problem-solving abilities have been employed in analog modality with plastic gear materials and wooden building blocks to test four- to seven-year-olds (Reuter & Leuchter, 2019, 2022; Weber et al., 2020). Likewise, the tasks assessing domain-specific knowledge have previously been employed using paper-and-pencil in the same age group

(Reuter & Leuchter, 2021; Weber & Leuchter, 2020). Since these analog tasks led to a great deal of effort and a high susceptibility to errors during implementation and video coding, one central aim of this dissertation was to adapt these tasks for tablet-based use. The tasks assessing the core executive functions (inhibition: Go/No-go task; working memory: Corsi blocks backward task; cognitive flexibility: *Flexible Item Selection* task) were developed based on previous studies using these tasks in the target age group (Gathercole et al., 2004; Jacques & Zelazo, 2001; Johann et al., 2020; St. John et al., 2019). All three executive function tasks were implemented with child-friendly stimuli and embedded in a fairytale-like cover story (cf. Johann & Karbach, 2018).

Table 1:

Tasks that were newly developed in the digital modality and employed in the dissertation studies

Task	Construct	Employed in Studies
Carousel	Science problem-solving	1 – 3
Propeller	Science problem-solving	1 – 3
Stabilization	Science problem-solving	1 – 3
Turning direction	Domain-specific knowledge	2
Turning speed	Domain-specific knowledge	2
Center-of-mass	Domain-specific knowledge	2
Go/No-go	Inhibition	3
Corsi blocks backward	Working memory	3
Flexible item selection	Cognitive flexibility	3

Note: All tasks were employed on tablets.

1.4.1 Requirements and structures of the science problem-solving tasks

Since this dissertation particularly emphasizes science problem-solving, the Carousel, Propeller, and Stabilization task structures will be described in more detail. Descriptions of the other tasks can be found in the methods sections of the original manuscripts of Study 2 (domain-specific knowledge) and Study 3 (executive functions).

Carousel task. The task goal is to connect two fixed gears on a gear-board using other movable gears so that both fixed gears turn in the same direction when one of them is turned. Participants have three minutes to complete the task and 14 movable gears of four different sizes are available. The domain-specific knowledge underlying this task is that adjacently connected gears turn in opposite directions.

Propeller task. The task goal is to attach two propellers to movable gears so that one of them turns as fast as possible and the other one as slow as possible. Both propellers need to be driven by a fixed gear on a gear-board and should not touch each other while turning. Participants have three minutes to complete the task and 14 movable gears of four different sizes, as well as two propellers are available. The domain-specific knowledge underlying this task is that a larger gear turns slower than a smaller gear.

Stabilization task. The task goal is to stabilize eight instable building block constructions by adding one additional building block. Up to three attempts can be used per block construction. The domain-specific knowledge underlying this task is that a building block construction is stable when the center-of-mass is supported.

Problem structures. Since all three science problem-solving tasks provide clearly defined goal states and pathways that guarantee to achieve those, they can be considered as well-defined. Consequently, the tasks provide objective ways to assess problem-solving performance (scoring systems are described in the original manuscript of Study 1). Moreover, the gear problem-solving tasks (i.e., Carousel and Propeller) offer opportunities to investigate the problem-solving process, by analyzing the way participants interact with the gears and propellers (see section 1.4.2).

The Carousel task can be considered to assess complex problem-solving because the spatial and functional properties of the task-relevant objects are interdependent (i.e., a gear turns left when an adjacently located gear turns right). The problem complexity in the Propeller and the Stabilization tasks is moderate because the spatial object properties are interdependent (e.g., a building block might have a particular orientation because it is stacked on another building block) but the functional object properties are independent (e.g., a gear's turning speed depends on its size and on the extrinsic movement impulse, but not on properties of adjacent gears).

All three tasks might further represent insight problems because testing intermediate solutions might provide participants with insights about the underlying domain-specific knowledge (e.g., turning adjacent gears and recognizing that they turn in opposite directions). However, insights are not essential for a good task performance because the participant might possess the domain-specific knowledge prior to task processing or successfully solve the tasks without this knowledge.

Moreover, there are causal associations between the task-relevant objects (e.g., an indirectly driven gear turns in a particular direction because its adjacent gear is turned in the opposite direction). Consequently, all three science problem-solving tasks may be considered as different kinds of reasoning tasks depending on the participant's prior domain-specific knowledge (cf. Bartley et al., 2018; Leuchter & Hardy, 2021): A participant who does not possess the domain-specific knowledge might acquire this knowledge inductively by testing intermediate solutions (e.g., by turning connected gear chains in the Carousel task and observe that adjacent gears turn in opposite directions) and, subsequently, apply it to solve the problems. A participant who possesses the domain-specific knowledge might deductively apply this knowledge to solve the problems (e.g., use the rule knowledge that adjacent gears turn in opposite directions to infer that two gears turn in the same direction when an odd number of other gears connects them). Irrespective of prior domain-specific knowledge, visuospatial relational reasoning might be required because spatial interdependencies between gears, propellers, and building blocks play a crucial role in each of the three tasks.

1.4.2 Implementation and data tracking

All tasks were programmed in the Unity engine (version 2020.3.17f1) using C# programming language and deployed in an integrative test software including auditive task instructions. The digital science problem-solving tasks exactly mimicked setup and functionalities of the analog materials true-to-scale, including size, color, shape, and distances of gears, propellers, building blocks, and the gear-board.

The tracking data that the test software automatically pre-processes, codes and stores, includes the solution quality as well as all actions a participant conducts in the science problem-solving tasks. Each data entity comprises the following information: The type and size of the manipulated object (i.e., gear, propeller, or building block), the type of action (i.e., a turning or a displacement), the time that the action was carried out, and the direction in which the object was moved (i.e., an object is moved onto/off the gear-board or block construction). Furthermore, for each task run, the total processing time and the total number of turnings and displacements are available in the data. The tracking data in the domain-specific knowledge and executive function tasks include response accuracy, reaction time, number of response selection switches, and type of error for each item.

2 Summary of Research Goals

The major objectives of this dissertation were to validate newly developed digital instruments assessing children's problem-solving abilities (Study 1) and to investigate to which extent domain-specific knowledge and the domain-general abilities intelligence and executive functions determine science problem-solving performance in elementary school children (Studies 2 and 3). Consequently, the aims were:

- Study 1: Testing the validity of new digital problem-solving tasks with gear constructions as compared to their traditional analog counterparts;

- Study 2: Assessing whether science problem-solving performance is more strongly affected by domain-specific knowledge of associated science concepts or by domain-general cognitive abilities;

- Study 3: Investigating whether the core executive functions inhibition, working memory, and cognitive flexibility individually contribute to science problem-solving performance.

3 Summary of the Empirical Studies

The following section provides short structured summaries of the three studies conducted within the framework of the present dissertation. Please see chapter 7 for the published versions of the full-length manuscripts.

3.1 Summary of Study 1

Schäfer, J., Reuter, T., Leuchter, M., & Karbach, J. (2024). Validation of new tablet-based problem-solving tasks in primary school students. *PLoS ONE*, *19*(8), e0309718. https://doi.org/10.1371/journal.pone.0309718

Background: Problem-solving abilities are strongly associated with positive educational outcomes and other cognitive abilities, such as language and reasoning. Nevertheless, empirical evidence on cognitive correlates of problem-solving performance in childhood is limited, in part because of a lack of valid and age-appropriate instruments assessing problem-solving performance. Most of such assessment instruments are based on analog tasks with play materials. However, given that analog tasks do not enable automated data tracking, evaluating performance scores requires time-consuming and error-prone coding. Thus, we developed and validated new tablet-based versions of existing analog tasks assessing problem-solving performance with gear construction tasks.

Methods: 215 children (six to eight years of age) performed the problem-solving tasks in both modalities (analog and digital). Additionally, participants performed three tasks assessing language, reasoning, and problem-solving with another content (stabilization) to assess related cognitive abilities (validation measures). Validity was assessed on different levels, namely by computing correlations between task modalities, testing for scalar measurement invariance across modalities, assessing the predictive values of the validation measures, and comparing the size of these predictive values.

Results: Performances on both task modalities were intercorrelated and also correlated with the validation measures, showing convergent validity. Structural equation modelling showed scalar measurement invariance across task modalities (represented as two distinct latent factors), supporting the validity of the digital tasks. Moreover, all validation measures had a significant predictive value for the performance on both modalities, except that reasoning did not significantly predict analog task performance. Constraining path coefficients showed that the size of these predictive values did not differ between modalities for language and problem-solving performance, but for reasoning.

Discussion: Overall, the analyses clearly confirm the validity of the new digital tasks. Thus, we conclude that the analog and the digital tasks draw on similar cognitive abilities and may be used interchangeably. The traditional analog tasks are successfully developed and validated for digital application, which is an important contribution to problem-solving research in childhood. The new digital tasks will provide substantial advantages for future studies, such as a standardized and reliable procedure, as well as efficient implementation and data evaluation. Moreover, the tablet-based instruments simplify the application in the field, e.g., in kindergartens or schools.

3.2 Summary of Study 2

Schäfer, J., Reuter, T., Karbach, J., & Leuchter, M. (2024). Domain-specific knowledge and domain-general abilities in children's science problem-solving. *British Journal of Educational Psychology, 94*(2), 346–366. https://doi.org/10.1111/bjep.12649

Background: Given the major importance of problem-solving abilities for subsequent achievement, it is of high interest for developmental psychologists to identify its cognitive correlates in early and middle childhood. Children's problem-solving performance in the realm of STEM, e.g., in science problem-solving, can be an essential indicator for pursuing technical professional careers. Previous research on problem-solving in childhood has rather focused on the problem-solving process (e.g., investigating strategies and problem-solving phases) than on the underlying cognitive processes and correlates of problem-solving performance. In particular, it was rarely investigated whether problem-solving performance rather relies on rule knowledge, that is specific to the problem content (domain-specific knowledge) or on cognitive abilities that may be applied across different problem domains (domain-general abilities). Consequently, the aim of Study 2 was to figure out whether science problem-solving performance primarily relies on domain-specific knowledge, on domain-general abilities, or on both.

Methods: 215 six- to eight-year-old children participated in this study. They completed three tasks measuring domain-specific knowledge on building block statics and gear turning mechanisms (e.g., "In which direction will this gear turn?") and three corresponding science problem-solving tasks (e.g., "Make these gears turn in the same direction"). Moreover, the children performed a language and a reasoning task as domain-general measures and proxies for intelligence. The analyses were based on the same sample as in Study 1.

Results: Correlational analyses, regression analyses, and structural equation modelling exhibited that, while the associations between science problem-solving performance across different domains were considerable, there were only small or no effects of domain-specific knowledge on science problem-solving performance. In contrast, the contribution of domain-general abilities (language and reasoning) to science problem-solving performance was significant and much stronger than the contribution of domain-specific knowledge. Moreover, age had a positive effect on science problem-solving performance.

Discussion: The findings of this study suggest that science problem-solving performance in six- to eight-year-old children strongly relies on domain-general abilities. This is inconsistent with previous studies arguing that children's problem-solving mainly relies on domain-specific knowledge but supports the assumption that there are cognitive resources being shared across problem domains. The positive age effects suggest that science problem-solving abilities improve during childhood. Future studies should consider a broader set of domain-general abilities, e.g., by including executive functions.

3.3 Summary of Study 3

Schäfer, J., Reuter, T., Leuchter, M., & Karbach, J. (2024). Executive functions and problem-solving – the contribution of inhibition, working memory, and cognitive flexibility to science problem-solving performance in elementary school students. *Journal of Experimental Child Psychology, 244*, 105962. https://doi.org/10.1016/j.jecp.2024.105962

Background: In light of the findings in Study 2, indicating that domain-general abilities are essential for children's science problem-solving, we conducted Study 3 focusing on executive functions as domain-general abilities. Previous research has shown that working memory, as one of three core executive functions, can contribute to successful problem-solving in pre- and elementary school children. However, since most studies did not simultaneously assess different aspects of executive functions, the role of both other core functions (inhibition and cognitive flexibility) is rather unclear. Consequently, Study 3 aimed to investigate the individual contribution of inhibition, working memory, and cognitive flexibility to science problem-solving performance in elementary school children.

Methods: 476 children from first and second grades (M_{age}=7.44 years) participated in Study 3. In the first experimental session, they performed a Go/No-go task (inhibition), a Corsi blocks backward task (working memory), and a flexible item selection task (cognitive flexibility) to assess the core executive functions. In the second session, they completed three science problem-solving tasks, including two gear turning tasks and one stabilization task. All tasks were completed tablet-based with the test software developed in the framework of Study 1.

Results: Structural equation modeling, including one latent factor for science problem-solving and one latent factor for each core executive function, yielded an excellent model fit. Regression paths from the executive function factors to the science problem-solving factor showed that working memory and cognitive flexibility contributed to science problem-solving performance, while inhibition did not. Equality constraining of regression paths revealed that the contribution of cognitive flexibility was significantly stronger than the contribution of working memory.

Discussion: The results suggest that executive functions play a significant role for science problem-solving in childhood. We conclude that working memory enables children to maintain task requirements and dynamic relations between task-relevant objects (i.e., gears, building blocks, and propellers), and cognitive flexibility supports the ability to switch between different problem-solving phases and dynamically changing problem states. Inhibitory processes may have a greater impact in tasks involving a higher degree of interference and in clinical populations exhibiting attentional deficits. Since working memory and cognitive flexibility have been shown to be essential components of successful science problem-solving in elementary school children, they should be considered in future research on science education.

4 General Discussion

The major aims of this dissertation were to develop test instruments for assessing problem-solving abilities and executive functions, and to identify the cognitive processes underlying science problem-solving in elementary school children.

The first major aim was accomplished through the development and validation of new tablet-based problem-solving and executive function tasks for pre- and elementary school children. The new problem-solving tasks were validated in Study 1 by demonstrating measurement invariance and convergent validity between analog and digital task modalities in six- to eight-year-old children. The executive function tasks were applied in Study 3 and their full validation in pre- and elementary school children will be the subject of a forthcoming study. Both task batteries will be made available to interested researchers, enabling a standardized format for diagnostics and intervention studies investigating cognitive abilities in childhood.

With regard to the cognitive processes underlying science problem-solving in childhood, the findings of Studies 2 and 3 of this dissertation provided new insights into a complex but rarely investigated research field. The results of Study 2 demonstrated that science problem-solving performance was more closely associated with domain-general abilities, as proxies for intelligence, than with domain-specific knowledge. The effects of language suggest that participants might have used language-based strategies, such as self-directed speech (Gunzenhauser et al., 2019). The effects of reasoning suggest that participants might have used elaborate mental strategies, such as scientific and evidence-based thinking (cf. Leuchter & Hardy, 2021; Tan et al., 2023). Nevertheless, the statistical associations between domain-general abilities and science problem-solving performance were not strong enough to conclude that intelligence and problem-solving abilities represent identical constructs (Hambrick et al., 2020). It is evident that further measures of intelligence would need to be considered to verify this conclusion. Moreover, the influence of domain-general abilities on problem-solving performance varied between science problem-solving tasks, which might be caused by different problem structures. Precisely, the Carousel task was considered more complex than the Propeller and Stabilization tasks (see section 1.4.1). Previous studies suggesting that the influence of higher-order cognition increases as a function of problem complexity (cf. Funke, 2010; Greiff et al., 2016) may explain why reasoning was a considerably strong predictor in the Carousel task (β = .66**). Study 2 also revealed positive age effects on science problem-solving performance, which confirmed previous research suggesting that problem-solving becomes more planful and goal-directed during childhood (Injoque-Ricle et al., 2014; Keen, 2011).

Study 3 indicated that some of the core executive functions contribute to science problem-solving performance. Importantly, the individual contribution of executive functions to science problem-solving performance varied from no significance (inhibition: β = .06) to considerable effect sizes (cognitive flexibility: β = .51**). The lack of an association with inhibition can be explained by the absence of task-irrelevant distractor objects in the science problem-solving tasks (cf. Lee et al., 2009). The significant effects of working memory (β = .23*) are in line with previous research (Greiff et al., 2016) and suggest a high relevance of remembering previously employed strategies and functional object interdependencies (regarding turning dynamics and stability). The strong effects of cognitive flexibility suggest a major importance of the ability to effectively switch between dynamically changing task demands and between different problem-solving phases (cf. Molnár & Greiff, 2023).

Taken together, Studies 2 and 3 showed that domain-general abilities had an impact on elementary school children's science problem-solving performance that went beyond the impact of domain-specific knowledge. This pattern of results might be unexpected in the light of previous research asserting that solving well-defined problems predominantly depends on domain-specific knowledge and involves only a limited degree of higher-order cognition (Schraw et al., 1995; Tricot & Sweller, 2014). However, the findings are in line with previous research suggesting a strong association between higher-order cognition and complex problem-solving (Greiff et al., 2016; Wu & Molnár, 2022; Wüstenberg et al., 2012) and, thereby, provide evidence for the assumption that the new problem-solving tasks entail a considerable complexity (see section 1.4.1; Funke, 2010).

4.1 Connections Between the Studies

In Study 1, new tablet-based science problem-solving tasks were developed and subsequently validated. These new digital tasks also served to assess science problem-solving performance in Studies 2 and 3. Consequently, Study 1 provided a basis for the subsequent studies to build on. Nevertheless, the findings from Study 1 extended beyond the validation of the new tasks to insights regarding the cognitive correlates of science problem-solving in childhood. More precisely, the confirmatory factor analyses revealed that language, reasoning, and stabilization problem-solving performance each explained variance in performance on the new digital tasks. This finding was investigated more closely in Study 2 by comparing the influence of domain-general abilities (i.e., language and reasoning) to the influence of domain-specific knowledge (i.e., rule knowledge of science concepts). Comparative hierarchical regression and confirmatory factor analyses revealed that the impact of intelligence on science problem-solving performance was stronger than the impact of domain-specific knowledge. Thus, the finding that domain-general abilities were more relevant for science problem-solving motivated Study 3 to investigate which other domain-general abilities play a role for children's science problem-solving. Consequently, Study 3 assessed the individual contribution of executive functions to science problem-solving performance.

Apart from the traditional analog problem-solving tasks in Study 1, all empirical tasks in the three studies were conducted digitally. Moreover, the new digital science problem-solving tasks were included in all three studies. This approach offers two key benefits: Firstly, it ensures methodological consistency across all three studies, and secondly, it allows for a direct comparison of their findings. Additionally, the inclusion criterion for participants to be aged between six and eight years was consistently maintained to keep the results comparable.

4.2 Limitations

It can be argued that the science problem-solving tasks conducted in the three dissertation studies represent very specific aspects of problem-solving and that results may not be representative beyond the science domain. The phenomena that informed the research approaches (i.e., turning dynamics and statics) are pertinent within the science context but may not be transferable to other content domains. Therefore, it should be considered that the cognitive underpinnings of problem-solving identified in this dissertation require further studies to be confirmed for other content domains.

The gear problem-solving tasks are limited in terms of increasing their task complexity (e.g., by removing available gears or increasing the distance between fixed gears). Moreover, it is quite possible to create satisfactory solutions by pure guessing, especially in the Carousel task. Given that the average solution quality in the Carousel task performance of six- to eight-year-old participants was 3.06 out of 4.00 points (data of Study 1), ceiling effects can be expected in adolescents and adults performing this task. Since a problem by definition involves the absence of an obvious solution (Duncker, 1945), ceiling effects would render a problem-solving measure invalid. Consequently, the tasks are unlikely to detect problem-solving abilities and their development beyond late childhood.

All validation measures used in Study 1 assessed cognitive abilities (language, reasoning, and stabilization problem-solving), while no validation measure assessed motor abilities. This implies that the convergent validity between the analog and the digital task modality was not controlled on the level of motor coordination (Pfister et al., 2014). However, it is plausible that gripping and pinning a plastic gear requires other motor skills than sliding a finger on a tablet (cf. Guilbert et al., 2019). Since this dissertation focused on the cognitive aspects of problem-solving, this circumstance does not impair the findings, but future studies that aim to consider motor skills should take this into account.

Interpretations of the dissertation findings need to consider that the investigated cognitive abilities provide a large theoretical overlap (Diamond, 2013). It has, for instance, been debated whether executive functions are equivalent with intelligence (cf. Friedman et al., 2006) and whether problem-solving is a higher-order executive function (Diamond, 2013; Zelazo et al., 1997). Furthermore, the psychological tradition of postulating cognitive abilities as clearly distinguishable constructs is increasingly being questioned due to the recognition of complex interdependencies between constructs and task-specific biases (De Boeck et al., 2023). This should not impede the development of differentiated study designs, but should motivate researchers to employ consistent terminology and avoid overgeneralization.

The analyses in the three studies focused on problem-solving performance but did not include measures on the problem-solving process, that may provide information about the strategies employed and the problem-solving phases. Thus, the three studies made significant contributions to our understanding of how well children with specific cognitive abilities solved science problems. However, identifying the way they solved these problems requires somewhat subjective interpretations, which introduces some uncertainty. Fortunately, the test software tracked abundant process data that enable future studies to additionally address strategies and problem-solving phases, as will be described in the following section.

4.3 Perspectives for Future Research

Since the studies of this dissertation focused on problem-solving performance only, future studies should closer address the problem-solving process. The large amount of tracking data (see section 1.4.2) that was generated in the scope of this dissertation by totally 691 participants offers potentials for this purpose. Precisely, the tracking data provide opportunities to investigate the following four research suggestions in pre- and elementary school children:

1) The association between science problem-solving performance and the number of gear turnings and displacements should be analyzed. A positive association between these variables would indicate that a trial-and-error strategy is more conducive to achieving a good solution, while a negative association would suggest that a more planful mental strategy, involving only goal-directed actions, is more likely to result in a good solution (cf. Injoque-Ricle, 2014; Molnár & Greiff, 2023).

2) The rate of gear turnings and displacements may indicate the specific problem-solving phase a participant is engaged in at a given point during task processing. Previous studies have designed assignments of the participant's actions and his or her progress in a problem-solving task (e.g., Wu & Molnár, 2022). Table 2 provides an assignment between the participant's behavior, indicated by the number of displacements and turnings, and the current problem-solving phase. This assignment can be applied to the time-coded tracking data of the participant's behavior in the gear problem-solving tasks and is reasoned as follows: In phase a), the number of turnings is small because there is not yet an available construction to be tested, but since understanding the problem requires exploration (Eichmann et al., 2019), the number of displacements is moderate. In phase b), the number of turnings and displacements is small because planning activates mental processes that precede motor actions (Rowe et al., 2001). In phase c), the number of displacements is large in order to build the gear construction

in the previously planned way (strategy execution), while a moderate number of turnings could be used to gather domain-specific knowledge as part of the strategy (Greiff et al., 2016). In phase d), a large number of turnings is required to test how the gears turn (Kendall, 2015), while displacements are avoided because the solution should not be changed at this stage. Notably, Table 2 describes a heuristic suggestion that requires validation, for example through human counter-coding.

Table 2:

Assignment of behaviour and the current problem-solving phase in the gear problem-solving tasks (i.e., Carousel and Propeller)

Problem-solving phase	Number of turnings	Number of displacements
a) Understanding and representing the problem	Small	Moderate
b) Planning/Developing a Strategy	Small	Small
c) Executing a strategy	Moderate	Large
d) Evaluating and monitoring solutions	Large	Small

Note: The numbers of turnings and displacements of gears within a specified time interval indicate the problem-solving phase that the participant is undergoing during that time interval.

3) Applying the assignment of Table 2 to the collected data would identify how many switches between problem-solving phases (i.e., transitions) a participant executed throughout a task. Based on these data, it could be investigated whether the number of transitions has an impact on science problem-solving performance. Previous studies suggest that iterative redesigning constructions improves children's engineering problem-solving performance

(e.g., Lucas et al., 2014). This may indicate that a greater number of transitions is beneficial for science problem-solving performance.

4) The relationship between cognitive flexibility and the number of executed transitions warrants further investigation. Previous research suggests that the proficiency in transitioning is positively related to problem-solving performance (Molnár & Greiff, 2023). Along this line, it can be investigated whether there are interaction effects of cognitive flexibility and the number of transitions on science problem-solving performance. To illustrate, it is reasonable that participants with a high cognitive flexibility commit more transitions, which improves their performance, since they iteratively refine their solutions (Lucas et al., 2014). This research suggestion is particularly pertinent to closer specify the role of cognitive flexibility in children's science problem-solving, that has been shown to be essential in Study 3 of this dissertation.

Research suggestions 1) and 2) might provide first insights to identifying the strategies that participants employed in the gear problem-solving tasks. More elaborate strategy analyses could extend to children's proactive and reactive cognitive control by using neuroimaging methods (Braver, 2012; Czernochowski, 2015). Neurobehavioral response patterns may indicate whether children anticipate the effects of their own actions (e.g., turning a gear and anticipate that an adjacent gear will turn in the opposite direction; proactive control) or not (e.g., turning a gear without particular expectations; reactive control). This again may suggest whether children follow a predetermined plan or rather use a trial-and-error approach.

In addition to the tracking data, human raters created video-based observational coding of the participants' behavior in both the analog and digital gear problem-solving tasks conducted in Study 1. The rater-coded data include the same variables as the tracking data, as well as variables estimating participants' handling proficiency, strategies, and quality of intermediate solutions. These data allow to add a view on potential differences in the problem-solving process between analog and digital task modality. Even though Study 1 showed convergent validity between task modalities on the performance level, it is still possible that some measures on the process level differ. For instance, participants might conduct more actions in the digital tasks due to a motorically simpler handling (i.e., swiping) compared to the analog tasks (i.e., pull out a gear and insert it into the board).

It has shown that the problem-solving process may differ between individual (as investigated in this dissertation) and collaborative problem-solving (i.e., solving a problem with several persons by sharing knowledge and skills; OECD, 2013). In collaborative problem-solving, the synergies of multiple perspectives might enhance problem-solving performance and might even foster learning gains on the individual level (Gauvain, 2018). Moreover, the interplay of problem-solving abilities and social skills makes collaborative problem-solving research particularly relevant for elementary school education (Hesse et al., 2015). The gear problem-solving tasks are well-suited to assess collaborative problem-solving because they integrate distinct aspects of domain-specific knowledge (i.e., turning direction and turning speed) that could be shared between participants.

Future research should consider a broader set of problem-solving tasks than those employed in this dissertation, including engineering design problems (Strimel et al., 2018), computational thinking (Montuori et al., 2024), and mathematical word problems (Swanson & Beebe-Frankenberger, 2004). Such an integrated STEM study design (Roberts et al., 2022) might verify whether the findings of this dissertation can be extended to other STEM domains. Along that study design, it would be beneficial to include more ill-defined problems, e.g., in a Microworld paradigm (Greiff et al., 2016), that still allow to comparably analyze the impact of cognitive abilities on problem-solving performance. Furthermore, assessing different aspects of reasoning (i.e., inductive, deductive, abductive, and visuospatial relational reasoning) and self-directed speech individually could closer verify the suggested structures of the investigated problems (see section 1.4.1; Bartley et al., 2018; Gunzenhauser et al., 2019; Leuchter & Hardy, 2021).

It has been suggested that the exploration of the problem space is an essential component of children's problem-solving, which provides a foundation for planning and strategy development (Appleton, 1995; Eichmann et al., 2019). Since STEM tasks typically require mentally representing task-relevant objects (Möhring et al., 2021; Ramey & Uttal, 2017), it is promising to conduct eye-tracking studies to investigate children's exploration of functional (e.g., turning dynamics) and spatial (e.g., size) object properties. Furthermore, eye-tracking data might verify whether the employed science problem-solving tasks are insight problems (see section 1.1.1; Gaschler et al., 2020) by identifying whether a child changes his or her strategy after observing particular events.

Given that this dissertation has found an important role of domain-general abilities for science problem-solving, it is promising to conduct cognitive training interventions (Karbach & Kray, 2009) to identify potential transfer effects to science problem-solving. There is evidence that cognitive trainings, focusing executive functions, may transfer to reading abilities in elementary school children (Johann & Karbach, 2020). Moreover, previous research showed transfer effects of executive function training to computational thinking in first-graders (Arfé et al., 2020). However, a study of Fyfe and Borriello (2024) failed to find transfer effects of domain-general abilities training on mathematics knowledge. Respective findings in the science domain are sparse.

5 Conclusion

In the framework of this dissertation, a number of new digital tasks for the empirical assessment of problem-solving abilities and executive functions in pre- and elementary school children were developed. In addition to this, two empirical studies were conducted with six- to eight-year-old children to investigate the impact of domain-specific knowledge, language, reasoning, inhibition, working memory, and cognitive flexibility on their science problem-solving performance. The results indicated that higher-order cognitive abilities, particularly language, reasoning, working memory, and cognitive flexibility had an essential impact on science problem-solving performance that was considerably larger than the impact of domain-specific knowledge. The findings are contrary to studies suggesting that (particularly well-defined) problem-solving relies on domain-specific knowledge and processes (Chi et al., 1981; Schraw et al., 1995) but provide evidence for problem-solving as a complex construct composed of higher-order cognitive abilities (Wang & Chiew, 2010). The effect sizes varied significantly across cognitive abilities and science problem-solving tasks, which might motivate future studies. Additionally, the substantial amount of generated but unused tracking data merits further investigation of the problem-solving process.

Traditional educational approaches have emphasized the acquisition of domain-specific content knowledge in a subject area as a key to learning progress (Block & Anderson, 1975; McKeachie, 1999). However, this dissertation suggests that higher-order cognitive abilities play a more essential role for how children solve science problems. This insight suggests that fostering these cognitive abilities may be more beneficial than merely focusing on teaching content knowledge. By prioritizing the development of language, reasoning and executive functions, educators could improve children's overall problem-solving abilities, potentially enhancing learning outcomes in science and other subjects.

6 References

Adams, E. J., Nguyen, A. T., & Cowan, N. (2018). Theories of working memory: Differences in definition, degree of modularity, role of attention, and purpose. *Language, speech, and hearing services in schools*, *49*(3), 340–355. https://doi.org/10.1044/2018_LSHSS-17-0114

Alderson-Day, B., & Fernyhough, C. (2015). Inner speech: Development, cognitive functions, phenomenology, and neurobiology. *Psychological Bulletin, 141*(5), 931–965. https://doi.org/10.1037/bul0000021

Allaire, J. C., & Marsiske, M. (2002). Well- and ill-defined measures of everyday cognition: Relationship to older adults' intellectual ability and functional status. *Psychology and Aging, 17*(1), 101–115. https://doi.org/10.1037/0882-7974.17.1.101

Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separable? *Child Development*, *77*(6), 1698–1716. https://doi.org/10.1111/j.1467-8624.2006.00968.x

Appleton, K (1995). Problem solving in science lessons: How students explore the problem space. *Research in Science Education*, *25*, 383–393. https://doi.org/10.1007/BF02357384

Arfé, B., Vardanega, T., & Ronconi, L. (2020). The effects of coding on children's planning and inhibition skills. *Computers & Education*, *148*, 103807. https://doi.org/10.1016/j.compedu.2020.103807

Astuti, N. H., Rusilowati, A., & Subali, B. (2021). STEM-based learning analysis to improve students' problem solving abilities in science subject: A literature review. *Journal of Innovative Science Education, 10*(1), 79–86. https://doi.org/10.15294/JISE.V9I2.38505

Avcil, M., & Artemenko, C. (2023). *Development of arithmetic across the lifespan: A registered report.* OSF. https://doi.org/10.17605/OSF.IO/9TVXH

Baddeley, A. D., & Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology, 8*(4), 485–493. https://doi.org/10.1037/0894-4105.8.4.485

Baldo, J. V., Dronkers, N. F., Wilkins, D., Ludy, C., Raskin, P., & Kim, J. (2005). Is problem solving dependent on language? *Brain and Language*, *92*(3), 240–250. https://doi.org/10.1016/j.bandl.2004.06.103

Bari, A., & Robbins, T. W. (2013). Inhibition and impulsivity: Behavioral and neural basis of response control. *Progress in Neurobiology*, *108*, 44–79. https://doi.org/10.1016/j.pneurobio.2013.06.005

Bartley, J. E., Boeving, E. R., Riedel, M. C., Bottenhorn, K. L., Salo, T., Eickhoff, S. B., Brewe, E., Sutherland, M. T., & Laird, A. R. (2018). Meta-analytic evidence for a core problem solving network across multiple representational domains. *Neuroscience and Biobehavioral Reviews*, *92*, 318–337. https://doi.org/10.1016/j.neubiorev.2018.06.009

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, *81*(6), 1641–1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Bignardi, G., Dalmaijer, E. S., Anwyl-Irvine, A., & Astle, D. E. (2021). Collecting big data with small screens: Group tests of children's cognition with touchscreen tablets are reliable and valid. *Behavior Research Methods*, *53*(4), 1515–1529. https://doi.org/10.3758/s13428-020-01503-3

Björngrim, S., Van den Hurk, W., Betancort, M., Machado, A., & Lindau, M. (2019). Comparing traditional and digitized cognitive tests used in standard clinical evaluation - A study of the digital application Minnemera. *Frontiers in Psychology*, *10*, 2327. https://doi.org/10.3389/fpsyg.2019.02327

Block, J. H., & Anderson, L. W. (1975). *Mastery learning in classroom instruction*. MacMillan.

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences*, *16*(2), 106–113. https://doi.org/10.1016/j.tics.2011.12.010

Brown, R. E. (2016). Hebb and Cattell: The genesis of the theory of fluid and crystallized intelligence. *Frontiers in Human Neuroscience*, *10*, 606. https://doi.org/10.3389/fnhum.2016.00606

Buades-Sitjar, F., & Duñabeitia, J. A. (2022). Intelligence subcomponents and their relationship to general knowledge. *Journal of Cultural Cognitive Science*, *6*, 343–355. https://doi.org/10.1007/s41809-022-00113-z

Buades-Sitjar, F., Boada, R., Guasch, M., Ferré, P., Hinojosa, J. A., & Duñabeitia, J. A. (2022). The predictors of general knowledge: Data from a Spanish megastudy. *Behavior Research Methods*, *54*(2), 898–909. https://doi.org/10.3758/s13428-021-01669-4

Bühner, M., Kröner, S., & Ziegler, M. (2008). Working memory, visual-spatial-intelligence and their relationship to problem-solving. *Intelligence, 36*(6), 672–680. https://doi.org/10.1016/j.intell.2008.03.008

Buttelmann, F., & Karbach, J. (2017). Development and plasticity of cognitive flexibility in early and middle childhood. *Frontiers in Psychology*, *8*, 1040. https://doi.org/10.3389/fpsyg.2017.01040

Cassotti, M., Agogué, M., Camarda, A., Houdé, O., & Borst, G. (2016). Inhibitory control as a core process of creative problem solving and idea generation from childhood to adult hood. In B. Barbot (Ed.), *Perspectives on creativity development. New Directions for Child and Adolescent Development* (pp. 61–72). John Wiley & Sons, Inc. https://doi.org/10.1002/cad.20153

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin, 40*(3), 153–193. https://doi.org/10.1037/h0059973

Chan, T. H. (1989). A statistical analysis of the towers of hanoi problem. *International Journal of Computer Mathematics*, *28*(1–4), 57–65. https://doi.org/10.1080/00207168908803728

Chan, W. W. L., & Kwan, J. L. Y. (2021). Pathways to word problem solving: The mediating roles of schema construction and mathematical vocabulary. *Contemporary Educational Psychology*, *65*, 101963. https://doi.org/10.1016/j.cedpsych.2021.101963

Chen, X., Hertzog, C., Park, D. C. (2017). Cognitive predictors of everyday problem solving across the lifespan. *Gerontology*, *63*(4), 372–384. https://doi.org/10.1159/000459622

Chen, Y., Spagna, A., Wu, T., Kim, T. H., Wu, Q., Chen, C., Wu, Y., & Fan, J. (2019). Testing a cognitive control model of human intelligence. *Scientific Reports*, *9*(1), 2898. https://doi.org/10.1038/s41598-019-39685-2

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices *. *Cognitive Science*, *5*(2), 121–152. https://doi.org/10.1207/s15516709cog0502_2

Conijn, J. M., Smits, N., & Hartman, E. E. (2020). Determining at what age children provide sound self-reports: An illustration of the Validity-Index Approach. *Assessment*, *27*(7), 1604-1618. https://doi.org/10.1177/1073191119832655

Corsi, P. (1972). *Human memory and the medial temporal region of the brain* [Doctoral dissertation, McGill University]. eScholarship@McGill. https://escholarship.mcgill.ca/concern/theses/05741s554

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science*, *19*(1), 51–57. https://doi.org/10.1177/0963721409359277

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, *26*(2), 197–223. https://doi.org/10.1007/s10648-013-9246-y

Csapó, B., Ainley, J., Bennett, R. E., Latour, T., & Law, N. (2012). Technological issues for computer-based assessment. In P. Griffin, B. McGaw, & E. Care (Eds.), *Assessment and Teaching of 21st Century Skills* (pp. 143-230). Springer. https://doi.org/10.1007/978-94-007-2324-5_4

Czernochowski, D. (2015). ERPs dissociate proactive and reactive control: Evidence from a task-switching paradigm with informative and uninformative cues. *Cognitive, Affective & Behavioral Neuroscience*, *15*, 117–131. https://doi.org/10.3758/s13415-014-0302-y

De Boeck, P., Pek, J., Walton, K., Wegener, D. T., Turner, B. M., Andersen, B. L., Beauchaine, T. P., Lecavalier, L., Myung, J. I., & Petty, R. E. (2023). Questioning psychological constructs: Current issues and proposed changes. *Psychological Inquiry*, *34*(4), 239–257. https://doi.org/10.1080/1047840X.2023.2274429

Deák, G. O., & Wiseheart, M. (2015). Cognitive flexibility in young children: General or task-specific capacity? *Journal of Experimental Child Psychology*, *138*, 31–53. https://doi.org/10.1016/j.jecp.2015.04.003

Demetriou, A., Spanoudis, G., & Mouyi, A. (2011). Educating the developing mind: Towards an overarching paradigm. *Educational Psychology Review*, *23*(4), 601–663. https://doi.org/10.1007/s10648-011-9178-3

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*, 135-168. https://doi.org/10.1146/annurev-psych-113011-143750

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science, 333*(6045), 959-964. https://doi.org/10.1126/science.1204529

diSessa, A. A. (1988). Knowledge in pieces. In G. Forman & P. Pufall (Eds.), *Constructivism in the computer age*. Lawrence Erlbaum Publishers.

Dixon, R. A., McFall, G. P., Whitehead, B. P., & Dolcos, S. (2013). Cognitive development in adulthood and aging. In R. M. Lerner, M. A. Easterbrooks, J. Mistry, & I. B. Weiner (Eds.), *Handbook of psychology: Developmental psychology* (2nd ed., pp. 451–474). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118133880.hop206018

Dörner, D., & Funke, J. (2017). Complex problem solving: What it is and what it is not. *Frontiers in Psychology*, *8*, 1153. https://doi.org/10.3389/fpsyg.2017.01153

Duncan, J., Chylinski, D., Mitchell, D. J., & Bhandari, A. (2017). Complexity and compositionality in fluid intelligence. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(20), 5295–5299. https://doi.org/10.1073/pnas.1621147114

Duncker, K. (1945). On problem-solving (L. S. Lees, Trans.). *Psychological Monographs, 58*(5), i–113. https://doi.org/10.1037/h0093599

D'Esposito, M. (2008). Chapter 11 Working memory. *Handbook of Clinical Neurology*, *88*, 237–247. https://doi.org/10.1016/S0072-9752(07)88011-0

Eichmann, B., Goldhammer, F., Greiff, S., Pucite, L., & Naumann, J. (2019). The role of planning in complex problem solving. *Computers & Education*, *128*, 1–12. https://doi.org/10.1016/j.compedu.2018.08.004

Ellis, D. M., Ball, B. H., Kimpton, N., & Brewer, G. A. (2020). The role of working memory capacity in analytic and multiply-constrained problem-solving in demanding situations. *Quarterly Journal of Experimental Psychology*, *73*(6), 920-928. https://doi.org/10.1177/1747021820909703

English, L. D. (2023). Ways of thinking in STEM-based problem solving. *ZDM - Mathematics Education*, *55*, 1219-1230. https://doi.org/10.1007/s11858-023-01474-7

Enke, S., Gunzenhauser, C., Hepach, R., Karbach, J., & Saalbach, H. (2022). Differences in cognitive processing? The role of verbal processes and mental effort in bilingual and monolingual children's planning performance. *Journal of Experimental Child Psychology*, *213*, 105255. https://doi.org/10.1016/j.jecp.2021.105255

Farrell Pagulayan, K., Busch, R. M., Medina, K. L., Bartok, J. A., & Krikorian, R. (2006). Developmental normative data for the Corsi Block-tapping task. *Journal of Clinical and Experimental Neuropsychology*, *28*(6), 1043–1052. https://doi.org/10.1080/13803390500350977

Forsberg, A., Adams, E. J., & Cowan, N. (2021). Chapter one - The role of working memory in long-term learning: Implications for childhood development. In: K. D. Federmeier (Ed.), *Psychology of Learning and Motivation* (pp. 1–45). Academic Press. https://doi.org/10.1016/bs.plm.2021.02.001

Fragkiadaki, G., Fleer, M., & Rai, P. (2023). Science concept formation during infancy, toddlerhood, and early childhood: Developing a scientific motive over time. *Research in Science Education*, *53*(2), 275–294. https://doi.org/10.1007/s11165-022-10053-x

Frensch, P. A., & Funke, J. (1995). Definitions, traditions, and a general framework for understanding complex problem solving. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (1st ed., pp. 3–25). Psychology Press.

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*(2), 172–179. https://doi.org/10.1111/j.1467-9280.2006.01681.x

Frischkorn, G. T., Wilhelm, O., & Oberauer, K. (2022). Process-oriented intelligence research: A review from the cognitive perspective. *Intelligence*, *94*, 101681. https://doi.org/10.1016/j.intell.2022.101681

Funke, J. (2003). *Problemlösendes Denken* (1st ed.). Kohlhammer.

Funke, J. (2010). Complex problem solving: A case for complex cognition? *Cognitive Processing*, *11*(2), 133–142. https://doi.org/10.1007/s10339-009-0345-0

Funke, J. (2012). Complex problem solving. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 682–685). Springer US. https://doi.org/10.1007/978-1-4419-1428-6_685

Funke, J., Fischer, A., & Holt, D. V. (2018). Competencies for complexity: Problem solving in the twenty-first century. In E. Care., P. Griffin., & M. Wilson (Eds.), *Educational Assessment in an Information Age: Assessment and Teaching of 21st Century Skills: Research and Applications* (pp. 41–53). Springer International Publishing. https://doi.org/10.1007/978-3-319-65368-6_3

Fyfe, E. R., & Borriello, G. A. (2024). The (in)effectiveness of training domain-general skills to support early math knowledge. *Child Development Perspectives*, 00, 1-7. Advance online publication. https://doi.org/10.1111/cdep.12526

Gao, Y., Zhai, X., Bulut, O., Cui, Y., & Sun, X. (2022). Examining humans' problem-solving styles in technology-rich environments using log file data. *Journal of Intelligence*, *10*(3), 38. https://doi.org/10.3390/jintelligence10030038

Gaschler, R. (2020). Komplexe Kognition. In T. Strobach (Ed.), *Kognitive Psychologie* (pp. 140–163). Kohlhammer. https://doi.org/10.17433/978-3-17-032662-0

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*(2), 177–190. https://doi.org/10.1037/0012-1649.40.2.177

Gauvain, M. (2018). Collaborative problem solving: Social and developmental considerations. *Psychological Science in the Public Interest*, *19*(2), 53–58. https://doi.org/10.1177/1529100618813370

Germine, L., Reinecke, K., & Chaytor, N. S. (2019). Digital neuropsychology: Challenges and opportunities at the intersection of science and software. *The Clinical Neuropsychologist*, *33*(2), 271–286. https://doi.org/10.1080/13854046.2018.1535662

Gold, Z. S., Elicker, J., Kellerman, A. M., Christ, S., Mishra, A. A., & Howe, N. (2021). Engineering play, mathematics, and spatial skills in children with and without disabilities. *Early Education and Development*, *32*(1), 49–65. https://doi.org/10.1080/10409289.2019.1709382

Greiff, S., & Fischer, A. (2013). Measuring complex problem solving: An educational application of psychological theories. *Journal for Educational Research Online, 5*, 34-53. https://hdl.handle.net/10993/3176

Greiff, S., Fischer, A., Stadler, M., & Wüstenberg, S. (2015a). Assessing complex problem-solving skills with multiple complex systems. *Thinking & Reasoning*, *21*(3), 356–382. https://doi.org/10.1080/13546783.2014.989263

Greiff, S., Krkovic, K., & Hautamäki, J. (2016). The prediction of problem-solving assessed via microworlds. *European Journal of Psychological Assessment*, *32*(4), 298–306. https://doi.org/10.1027/1015-5759/a000263

Greiff, S., Wüstenberg, S., & Avvisati, F. (2015b). Computer-generated log-file analyses as a window into students' minds? A showcase study based on the PISA 2012 assessment of problem solving. *Computers & Education*, *91*, 92–105. https://doi.org/10.1016/j.compedu.2015.10.018

Grubbs, M. E., Strimel, G. J., & Kim, E. (2018). Examining design cognition coding schemes for P-12 engineering/technology education. *International Journal of Technology and Design Education, 28*, 899–920. https://doi.org/10.1007/s10798-017-9427-y

Guilbert, J., Alamargot, D., & Morin, M.-F. (2019). Handwriting on a tablet screen: Role of visual and proprioceptive feedback in the control of movement by children and adults. *Human Movement Science*, *65.* https://doi.org/10.1016/j.humov.2018.09.001

Gunzenhauser, C., Karbach, J., & Saalbach, H. (2019). Function of verbal strategies in monolingual vs. bilingual students’ planning performance: An experimental approach. *Cognitive Development*, *50*, 1–12. https://doi.org/10.1016/j.cogdev.2019.01.003

Halfin, H. H. (1973). *Technology: A process approach* (Publication No. 7323867) [Doctoral dissertation, West Virginia University]. West Virginia University ProQuest Dissertations & Theses.

Hambrick, D. Z., Burgoyne, A. P., & Altmann, E. M. (2020). Problem-solving and intelligence. In R. J. Sternberg (Ed.), *The Cambridge Handbook of Intelligence* (pp. 553–579). Cambridge University Press. https://doi.org/10.1017/9781108770422.024

Harada, C. N., Natelson Love, M. C., & Triebel, K. L. (2013). Normal cognitive aging. *Clinics in Geriatric Medicine*, *29*(4), 737–752. https://doi.org/10.1016/j.cger.2013.07.002

Heaton, R. K., Akshoomoff, N., Tulsky, D., Mungas, D., Weintraub, S., Dikmen, S., Beaumont, J., Casaletto, K. B., Conway, K., Slotkin, J., & Gershon, R. (2014). Reliability and validity of composite scores from the NIH Toolbox Cognition Battery in adults. *Journal of the International Neuropsychological Society*, *20*(6), 588–598. https://doi.org/10.1017/S1355617714000241

Hesse, F., Care, E., Buder, J., Sassenberg, K., & Griffin, P. (2015). A framework for teachable collaborative problem solving skills. In P. Griffin & E. Care (Eds.), *Educational Assessment in an Information Age: Assessment and Teaching of 21st Century Skills: Methods and Approach* (pp. 37-56). Springer. https://doi.org/10.1007/978-94-017-9395-7_2

Hidayati, R. M., & Wagiran, W. (2020). Implementation of problem-based learning to improve problem-solving skills in vocational high school. *Jurnal Pendidikan Vokasi*, *10*(2), 177—187. https://doi.org/10.21831/jpv.v10i2.31210

Hodgkiss, A., Gilligan, K. A., Tolmie, A. K., Thomas, M. S. C., & Farran, E. K. (2018). Spatial cognition and science achievement: The contribution of intrinsic and extrinsic spatial skills from 7 to 11 years. *British Journal of Educational Psychology*, *88*(4), 675–697. https://doi.org/10.1111/bjep.12211

Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences*, *16*(3), 174–180. https://doi.org/10.1016/j.tics.2012.01.006

Hollenstein, L., Thurnheer, S., & Vogt, F. (2022). Problem solving and digital transformation: Acquiring skills through pretend play in kindergarten. *Education Sciences*, *12*(2), 92. https://doi.org/10.3390/educsci12020092

Holmes, J., & Adams, J. W. (2006). Working memory and children's mathematical skills: Implications for mathematical development and mathematics curricula. *Educational Psychology*, *26*(3), 339–366. https://doi.org/10.1080/01443410500341056

Hurlstone, M. J., & Hitch, G. J. (2018). How is the serial order of a visual sequence represented? Insights from transposition latencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(2), 167–192. https://doi.org/10.1037/xlm0000440

Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2014). Memory for serial order across domains: An overview of the literature and directions for future research. *Psychological Bulletin*, *140*(2), 339–373. https://doi.org/10.1037/a0034221

Injoque-Ricle, I., Barreyro, J. P., Calero, A., & Burin, D. I. (2014). Tower of London: Planning development in children from 6 to 13 years of age. *The Spanish Journal of Psychology*, *17*, E77. https://doi.org/10.1017/sjp.2014.83

Jacques, S., & Zelazo, P. D. (2001). The Flexible Item Selection Task (FIST): A measure of executive function in preschoolers. *Developmental Neuropsychology*, *20*(3), 573–591. https://doi.org/10.1207/S15326942DN2003_2

Johann, V. E., & Karbach, J. (2018). Validation of new online game-based executive function tasks for children. *Journal of Experimental Child Psychology*, *176*, 150–161. https://doi.org/10.1016/j.jecp.2018.07.009

Johann, V. E., & Karbach, J. (2020). Effects of game-based and standard executive control training on cognitive and academic abilities in elementary school children. *Developmental Science*, *23*(4), e12866. https://doi.org/10.1111/desc.12866

Johann, V. E., Könen, T., & Karbach, J. (2020). The unique contribution of working memory, inhibition, cognitive flexibility, and intelligence to reading comprehension and reading speed. *Child Neuropsychology*, *26*(3), 324–344. https://doi.org/10.1080/09297049.2019.1649381

Johnstone, S. J., Dimoska, A., Smith, J. L., Barry, R. J., Pleffer, C. B., Chiswick, D., & Clarke, A. R. (2007). The development of stop-signal and Go/Nogo response inhibition in children aged 7-12years: Performance and event-related potential indices.

International Journal of Psychophysiology, *63*(1), 25–38. https://doi.org/10.1016/j.ijpsycho.2006.07.001

Jonassen, D. H. (1997). Instructional design models for well-structured and III-structured problem-solving learning outcomes. *Educational Technology Research and Development, 45*, 65–94. https://doi.org/10.1007/BF02299613

Jonassen, D. H. (2013). First principles of learning. In J. M. Spector, B. B. Lockee, S. Smaldino, & M. Herring (Eds.), *Learning, Problem Solving, and Mindtools: Essays in Honor of David H. Jonassen* (1st ed.). Routledge.

Karbach, J., & Kray, J. (2009). How useful is executive control training? Age differences in near and far transfer of task-switching training. *Developmental Science*, *12*(6), 978–990. https://doi.org/10.1111/j.1467-7687.2009.00846.x

Keen, R. (2011). The development of problem solving in young children: A critical cognitive skill. *Annual Review of Psychology*, *62*, 1–21. https://doi.org/10.1146/annurev.psych.031809.130730

Keith Berg, W., & Byrd, D. L. (2002). The Tower of London spatial problem-solving task: Enhancing clinical and research implementation. *Journal of Clinical and Experimental Neuropsychology*, *24*(5), 586–604. https://doi.org/10.1076/jcen.24.5.586.1006

Kendall, A. L. M. (2015, June 14-17). *Evaluate-and-redesign tasks: Using interviews to investigate how elementary students iterate (fundamentals).* [Paper presentation]. 2015 ASEE Annual Conference & Exposition, Seattle, Washington, United States. https://doi.org/10.18260/p.24025

Kim, M., & Pegg, J. (2019). Case analysis of children's reasoning in problem-solving process. *International Journal of Science Education*, *41*(6), 739–758. https://doi.org/10.1080/09500693.2019.1579391

Kompa, N. A., & Mueller, J. L. (2022). Inner speech as a cognitive tool—or what is the point of talking to oneself? *Philosophical Psychology*, 1–24. https://doi.org/10.1080/09515089.2022.2112164

Korovkin, S., Vladimirov, I., Chistopolskaya, A., & Savinova, A. (2018). How working memory provides representational change during insight problem solving. *Frontiers in Psychology*, *9*, 1864. https://doi.org/10.3389/fpsyg.2018.01864

Kretzschmar, A., & Nebe, S. (2021). Working memory, fluid reasoning, and complex problem solving: Different results explained by the Brunswik symmetry. *Journal of Intelligence*, *9*(1), 5. https://doi.org/10.3390/jintelligence9010005

Kretzschmar, A., Neubert, J. C., Wüstenberg, S., & Greiff, S. (2016). Construct validity of complex problem solving: A comprehensive view on different facets of intelligence and school grades. *Intelligence*, *54*, 55–69. https://doi.org/10.1016/j.intell.2015.11.004

Krist, H. (2010). Development of intuitions about support beyond infancy. *Developmental Psychology*, *46*(1), 266–278. https://doi.org/10.1037/a0018040

Kröner, S., Plass, J. L., & Leutner, D. (2005). Intelligence assessment with computer simulations. *Intelligence*, *33*(4), 347–368. https://doi.org/10.1016/j.intell.2005.03.002

Lee, K., Ng, E. L., & Ng, S. F. (2009). The contributions of working memory and executive functioning to problem representation and solution generation in algebraic word problems. *Journal of Educational Psychology, 101*(2), 373–387. https://doi.org/10.1037/a0013843

Lehrer, R., & Schauble, L. (1998). Reasoning about structure and function: Children's conceptions of gears. *Journal of Research in Science Teaching, 35*(1), 3-25. https://doi.org/10.1002/(SICI)1098-2736(199801)35:1<3::AID-TEA2>3.0.CO;2-X

Leighton, J. P., & Sternberg, R. J. (2003). Reasoning and problem solving. In A. F. Healy & R. W. Proctor (Eds.), *Handbook of psychology: Experimental psychology* (4th ed., pp. 623–648). John Wiley & Sons, Inc. https://doi.org/10.1002/0471264385.wei0423

Leuchter, M., & Hardy, I. (2021). Kognitive Prozesse als Grundlage des wissenschaftlichen Denkens und Argumentierens im frühen und mittleren Kindesalter. *Unterrichtswissenschaft*, *49*(1), 17–30. https://doi.org/10.1007/s42010-021-00095-x

Lile Diamond, L. (2018). Problem solving in the early years. *Intervention in School and Clinic, 53*(4), 220–223. https://doi.org/10.1177/1053451217712957

Liu, X., Han, X., Wang, T., & Ren, X. (2024). An account of the relationship between critical thinking and fluid intelligence in considering executive functions. *Thinking Skills and Creativity*, *52*, 101538. https://doi.org/10.1016/j.tsc.2024.101538

Lucas, B., Hanson, J., & Claxton, G. (2014). *Thinking like an engineer: implications for the education system*. London: Royal Academy of Engineering. Available under:

https://www.stem.org.uk/system/files/elibrary-resources/2018/01/thinking-like-an-engineer-implications-summary.pdf

Lynch, C., Ashley, K. D., Pinkwart, N., & Aleven, V. (2009). Concepts, structures, and goals: Redefining ill-definedness. *International Journal of Artificial Intelligence in Education*, *19*(3), 253-266.

Marcus, O., Martins, E. C., Sassu, R., & Visu-Petra, L. (2020). On the importance of being flexible: early interrelations between affective flexibility, executive functions and anxiety symptoms in preschoolers. *Early Child Development and Care*, *192*(6), 914–931. https://doi.org/10.1080/03004430.2020.1816995

Mayer, D., Sodian, B., Koerber, S., & Schwippert, K. (2014). Scientific reasoning in elementary school children: Assessment and relations with cognitive abilities. *Learning and Instruction*, *29*, 43–55. https://doi.org/10.1016/j.learninstruc.2013.07.005

Mayer, R. E., & Wittrock, M. C. (2006). Problem solving. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (pp. 287–303). Lawrence Erlbaum Associates Publishers.

Mayer, R. E. (1992). [Review of the book *Complex Problem Solving: Principles and Mechanisms*, by R. J. Sternberg & P. A. Frensch]. *The American Journal of Psychology*, *105*(3), 501–504. https://doi.org/10.2307/1423205

Mayer, R. E. (1998). Cognitive, metacognitive, and motivational aspects of problem solving. *Instructional Science*, *26*(1), 49–63. https://doi.org/10.1023/A:1003088013286

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 31–48). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.004

McKeachie, W. J. (1999). *McKeachies's teaching tips: Strategies, research and theory for college and university teachers* (10th ed.). Houghton Mifflin & Co.

Mehadi Rahman, M. (2019). 21st century skill "Problem solving": Defining the concept. *Asian Journal of Interdisciplinary Research*, *2*(1), 64–74. https://doi.org/10.34256/ajir1917

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Möhring, W., Ribner, A. D., Segerer, R., Libertus, M. E., Kahl, T., Troesch, L. M., & Grob, A. (2021). Developmental trajectories of children's spatial skills: Influencing variables and associations with later mathematical thinking. *Learning and Instruction*, *75*, 101515. https://doi.org/10.1016/j.learninstruc.2021.101515

Molnár, G., & Csapó, B. (2018). The efficacy and development of students' problem-solving strategies during compulsory schooling: Logfile analyses. *Frontiers in Psychology*, *9*, 302. https://doi.org/10.3389/fpsyg.2018.00302

Molnár, G., & Greiff, S. (2023). Understanding transitions in complex problem-solving: Why we succeed and where we fail. *Thinking Skills and Creativity*, *50*, 101408. https://doi.org/10.1016/j.tsc.2023.101408

Montuori, C., Gambarota, F., Altoé, G., & Arfé, B. (2024). The cognitive effects of computational thinking: A systematic review and meta-analytic study. *Computers & Education*, *210*, 104961. https://doi.org/10.1016/j.compedu.2023.104961

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

OECD. (2013). *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*. OECD Publishing. https://dx.doi.org/10.1787/9789264190511-en

OECD. (2014). *PISA 2012 Results: Creative Problem Solving (Volume V): Students' Skills in Tackling Real-Life Problems*. OECD Publishing. https://doi.org/10.1787/9789264208070-en

Ozonoff, S., Goodlin-Jones, B. L., & Solomon, M. (2005). Evidence-based assessment of autism spectrum disorders in children and adolescents. *Journal of Clinical Child & Adolescent Psychology*, *34*(3), 523–540. https://doi.org/10.1207/s15374424jccp3403_8

O'Brien, S., Mitchell, D. J., Duncan, J., & Holmes, J. (2023). Cognitive segmentation and fluid reasoning in childhood. *Quarterly Journal of Experimental Psychology*, *76*(6), 1431–1444. https://doi.org/10.1177/17470218221116054

Perkins, D. N., & Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher*, *18*(1), 16–25. https://doi.org/10.3102/0013189X018001016

Pfister, R., Janczyk, M., Gressmann, M., Fournier, L. R., & Kunde, W. (2014). Good vibrations? Vibrotactile self-stimulation reveals anticipation of body-related action effects in motor control. *Experimental Brain Research*, *232*(3), 847–854. https://doi.org/10.1007/s00221-013-3796-6

Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures* (A. Rosin Trans.). Viking.

Pólya, G. (1957). *How to solve it: A new aspect of mathematical method* (2nd ed.). Princeton University Press.

Priemer, B., Eilerts, K., Filler, A., Pinkwart, N., Rösken-Winter, B., Tiemann, R., & Upmeier zu Belzen, A. (2019). A framework to foster problem-solving in STEM and computing education. *Research in Science & Technological Education*, *38*(1), 105–130. https://doi.org/10.1080/02635143.2019.1600490

Ramey, K. E., & Uttal, D. H. (2017). Making sense of space: Distributed spatial sensemaking in a middle school summer engineering camp. *Journal of the Learning Sciences*, *26*(2), 277–319. https://doi.org/10.1080/10508406.2016.1277226

Reuter, T., & Leuchter, M. (2019). *Preschooler's concepts and problem solving in the domain of gears in the context of guided play* [Conference presentation]. 2019 18th biennial EARLI Conference, Aachen, Germany.

Reuter, T., & Leuchter, M. (2021). Children's concepts of gears and their promotion through play. *Journal of Research in Science Teaching*, *58*(1), 69–94. https://doi.org/10.1002/tea.21647

Reuter, T., & Leuchter, M. (2022). Examining kindergarten children's testing and optimising in the context of a gear engineering task. *European Journal of STEM Education*, *7*(1), 4. https://doi.org/10.20897/ejsteme/11827

Rey-Mermet, A., Gade, M., & Oberauer, K. (2018). Should we stop thinking about inhibition? Searching for individual and age differences in inhibition ability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(4), 501–526. https://doi.org/10.1037/xlm0000450

Rigas, G., & Brehmer, B. (1999). Mental processes in intelligence tests and dynamics decision making tasks. In P. Juslin & H. Montgomerry (Eds.), *Judgment and decision making: Neo-brunswikian and Process-tracing approaches* (1st ed., pp. 45–65). Psychology Press.

Roberts, M. J. (Ed.). (2007). *Integrating the mind: Domain general vs domain specific processes in higher cognition.* Psychology Press.

Roberts, T., Maiorca, C., Jackson, C., & Mohr-Schroeder, M. (2022). Integrated STEM as problem-solving practices. *Investigations in Mathematics Learning*, *14*(1), 1–13. https://doi.org/10.1080/19477503.2021.2024721

Rohrkemper, M. (1986). The functions of inner speech in elementary school students' problem-solving behavior. *American Educational Research Journal*, *23*(2), 303–313. https://doi.org/10.2307/1162961

Rowe, J. B., Owen, A. M., Johnsrude, I. S., & Passingham, R. E. (2001). Imaging the mental components of a planning task. *Neuropsychologia, 39*(3), 315-327. https://doi.org/10.1016/S0028-3932(00)00109-3

Salthouse, T. (2012). Consequences of age-related cognitive declines. *Annual Review of Psychology*, *63*, 201–226. https://doi.org/10.1146/annurev-psych-120710-100328

Sarathy, V. (2018). Real world problem-solving. *Frontiers in Human Neuroscience*, *12*, 261. https://doi.org/10.3389/fnhum.2018.00261

Schäfer, J., Reuter, T., Karbach, J., & Leuchter, M. (2024). Domain-specific knowledge and domain-general abilities in children's science problem-solving. *British Journal of Educational Psychology, 94*(2), 346–366. https://doi.org/10.1111/bjep.12649

Schäfer, J., Reuter, T., Leuchter, M., & Karbach, J. (2024). Executive functions and problem-solving – the contribution of inhibition, working memory, and cognitive flexibility to science problem-solving performance in elementary school students. *Journal of Experimental Child Psychology, 244*, 105962. https://doi.org/10.1016/j.jecp.2024.105962

Schäfer, J., Reuter, T., Leuchter, M., & Karbach, J. (2024). Validation of new tablet-based problem-solving tasks in primary school students. *PLoS ONE*, *19*(8), e0309718. https://doi.org/10.1371/journal.pone.0309718

Schmidt, J. C. (2011). What is a problem?: On problem-oriented interdisciplinarity. *Poiesis & Praxis*, *7*, 249–274. https://doi.org/10.1007/s10202-011-0091-0

Schoenfeld, A. H. (2016). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics (Reprint). *Journal of Education*, *196*(2), 1–38. https://doi.org/10.1177/002205741619600202

Schraw, G., Dunkle, M. E., & Bendixen, L. D. (1995). Cognitive processes in well-defined and ill-defined problem solving. *Applied Cognitive Psychology, 9*(6), 523–538. https://doi.org/10.1002/acp.2350090605

Schunn, C. D., McGregor, M. U., & Saner, L. D. (2005). Expertise in ill-defined problem-solving domains as effective strategy use. *Memory & Cognition, 33*(8), 1377–1387. https://doi.org/10.3758/BF03193370

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *298*(1089), 199–209. https://doi.org/10.1098/rstb.1982.0082

Shechter, T., Eden, S., & Spektor-Levy, O. (2021). Preschoolers' nascent engineering thinking during a construction task. *Journal of Cognitive Education and Psychology, 20*(2), 83–111. https://doi.org/10.1891/JCEP-D-20-00010

Smith, M. K., Jones, F. H. M., Gilbert, S. L., & Wieman, C. E. (2013). The classroom observation protocol for undergraduate STEM (COPUS): A new instrument to characterize university STEM classroom practices. *CBE Life Sciences Education*, *12*(4), 577-723. https://doi.org/10.1187/cbe.13-08-0154

Soltanlou, M., Artemenko, C., Dresler, T., Haeussinger, F. B., Fallgatter, A. J., Ehlis, A.-C., & Nuerk, H.-C. (2017). Increased arithmetic complexity is associated with domain-general but not domain-specific magnitude processing in children: A simultaneous fNIRS-EEG study. *Cognitive, Affective, & Behavioral Neuroscience*, *17*(4), 724–736. https://doi.org/10.3758/s13415-017-0508-x

Somsen, R. J. M. (2007). The development of attention regulation in the Wisconsin Card Sorting Task. *Developmental Science*, *10*(5), 664–680. https://doi.org/10.1111/j.1467-7687.2007.00613.x

Stadler, M., Becker, N., Gödker, M., Leutner, D., & Greiff, S. (2015). Complex problem solving and intelligence: A meta-analysis. *Intelligence*, *53*, 92–101. https://doi.org/10.1016/j.intell.2015.09.005

Sternberg, R. J. (2012). Intelligence. *Dialogues in clinical neuroscience*, *14*(1), 19–27. https://doi.org/10.31887/DCNS.2012.14.1/rsternberg

Stoffers, D., Berendse, H. W., Deijen, J. B., & Wolters, E. C. (2003). Deficits on Corsi's block-tapping task in early stage Parkinson's disease. *Parkinsonism & Related Disorders*, *10*(2), 107–111. https://doi.org/10.1016/s1353-8020(03)00106-8

Strimel, G. J., Bartholomew, S. R., Kim, E., & Zhang, L. (2018). An investigation of engineering design cognition and achievement in primary school. *Journal for STEM Education Research, 1*, 173–201. https://doi.org/10.1007/s41979-018-0008-0

St. John, A. M., Finch, K., & Tarullo, A. R. (2019). Socioeconomic status and neural processing of a go/no-go task in preschoolers: An assessment of the P3b. *Developmental Cognitive Neuroscience*, *38*, 100677. https://doi.org/10.1016/j.dcn.2019.100677

Swanson, H. L., & Beebe-Frankenberger, M. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology*, *96*(3), 471–491. https://doi.org/10.1037/0022-0663.96.3.471

Tan, A.-L., Ong, Y. S., Ng, Y. S., & Tan, J. H. J. (2023). STEM problem solving: Inquiry, concepts, and reasoning. *Science & Education*, *32*(2), 381–397. https://doi.org/10.1007/s11191-021-00310-2

Tager-Flusberg H., & Plesa-Skwerer, D. (2009). Chapter 81 - Assessment of intelligence. In W. B. Carey, A. C. Crocker, W. L. Coleman, E. R. Elias, & H. M. Feldman (Eds.), *Developmental-Behavioral Pediatrics* (4th ed., pp. 803-809). Elsevier Inc.

Thalmann, M., Souza, A. S., & Oberauer, K. (2019). How does chunking help working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(1), 37–55. https://doi.org/10.1037/xlm0000578

Titz, C., & Karbach, J. (2014). Working memory and executive functions: Effects of training on academic achievement. *Psychological Research*, *78*, 852–868. https://doi.org/10.1007/s00426-013-0537-1

Tönnsen, K.-C. (2021). The relevance of trial-and-error: Can trial-and-error be a sufficient learning method in technical problem-solving-contexts? *Techne Serien – Forskning i slöjdpedagogik Och slöjdvetenskap*, *28*(2), 303–312.

Träff, U., Skagerlund, K., Östergren, R., & Skagenholt, M. (2023). The importance of domain-specific number abilities and domain-general cognitive abilities for early arithmetic achievement and development. *British Journal of Educational Psychology*, *93*(3), 825–841. https://doi.org/10.1111/bjep.12599

Tricot, A., & Sweller, J. (2014). Domain-specific knowledge and why teaching generic skills does not work. *Educational Psychology Review, 26*, 265–283. https://doi.org/10.1007/s10648-013-9243-1

Van den Haak, M., De Jong, M., & Schellens, P. J. (2003). Retrospective vs. concurrent think-aloud protocols: Testing the usability of an online library catalogue. *Behaviour & Information Technology, 22*(5), 339–351. https://doi.org/10.1080/0044929031000

Vermeent, S., Spaltman, M., Van Elswijk, G., Miller, J. B., & Schmand, B. (2022). Philips IntelliSpace Cognition digital test battery: Equivalence and measurement invariance compared to traditional analog test versions. *The Clinical Neuropsychologist*, *36*(8), 2278–2299. https://doi.org/10.1080/13854046.2021.1974565

Viterbori, P., Traverso, L., & Usai, M. C. (2017). The role of executive function in arithmetic problem-solving processes: A study of third graders. *Journal of Cognition and Development*, *18*(5), 595–616. https://doi.org/10.1080/15248372.2017.1392307

Wallace, G. L., Peng, C. S., & Williams, D. (2017). Interfering with inner speech selectively disrupts problem solving and is linked with real-world executive functioning. *Journal of Speech, Language, and Hearing Research : JSLHR*, *60*(12), 3456–3460. https://doi.org/10.1044/2017_JSLHR-S-16-0376

Wang, Y., & Chiew, V. (2010). On the cognitive process of human problem solving. *Cognitive Systems Research*, *11*(1), 81–92. https://doi.org/10.1016/j.cogsys.2008.08.003

Weber, A. M., & Leuchter, M. (2020). Measuring preschool children's knowledge of the principle of static equilibrium in the context of building blocks: Validation of a test instrument. *British Journal of Educational Psychology*, *90*(S1), 50–74. https://doi.org/10.1111/bjep.12304

Weber, A. M., Reuter, T., & Leuchter, M. (2020). The impact of a construction play on 5- to 6-year-old children's reasoning about stability. *Frontiers in Psychology*, *11*, 1737. https://doi.org/10.3389/fpsyg.2020.01737

Wechsler, D. (2012). *Wechsler Preschool and Primary Scale of Intelligence* (4th ed.). Pearson.

Whitelock-Wainwright, A., Laan, N., Wen, D., & Gašević, D. (2020). Exploring student information problem solving behaviour using fine-grained concept map and search tool data. *Computers & Education*, *145*, 103731. https://doi.org/10.1016/j.compedu.2019.103731

Wicaksono, A. G. C., & Korom, E. (2022). Review of problem-solving measurement: An assessment developed in the Indonesian context. *Participatory Educational Research*, *9*(1), 116–136. https://doi.org/10.17275/per.22.7.9.1

Wicklein, R. C., & Rojewski, J. W. (1999). Toward a "Unified Curriculum Framework" for technology education. *Journal of Industrial Teacher Education*, *36*(4).

Wiebe, S. A., & Karbach, J. (2017). *Executive function: Development across the life span* (1st ed.). Routledge. https://doi.org/10.4324/9781315160719

Wilkening, F., & Cacchione, T. (2014). Children’s intuitive physics. In U. Goswami (Ed.), *Wiley-Blackwell handbooks of developmental psychology. The Wiley-Blackwell* *handbook of childhood cognitive development* (2nd ed., pp. 473–496). John Wiley & Sons, Inc.

Wu, H., & Molnár, G. (2022). Analysing complex problem-solving strategies from a cognitive perspective: The role of thinking skills. *Journal of Intelligence*, *10*(3), 46. https://doi.org/10.3390/jintelligence10030046

Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex problem solving — More than reasoning? *Intelligence*, *40*(1), 1–14. https://doi.org/10.1016/j.intell.2011.11.003

Yeniad, N., Malda, M., Mesman, J., Van IJzendoorn, M. H., & Pieper, S. (2013). Shifting ability predicts math and reading performance in children: A meta-analytical study. *Learning and Individual Differences*, *23*, 1–9. https://doi.org/10.1016/j.lindif.2012.10.004

Yeo, J., & Tan, S. C. (2014). Redesigning problem-based learning in the knowledge creation paradigm for school science learning. *Instructional Science*, *42*(5), 747–775. https://doi.org/10.1007/s11251-014-9317-6

Zamora, E. V., Vernucci, S., del Valle, M., Introzzi, I., & Richard's, M. M. (2020). Assessing cognitive inhibition in emotional and neutral contexts in children. *Educational and Developmental Psychologist*, *37*(1), 56–66. https://doi.org/10.1017/edp.2020.4

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, *1*(1), 297–301. https://doi.org/10.1038/nprot.2006.46

Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology, 1*(2), 198–226. https://doi.org/10.1037/1089-2680.1.2.198

Zook, N. A., Davalos, D. B., DeLosh, E. L., & Davis, H. P. (2004). Working memory, inhibition, and fluid intelligence as predictors of performance on Tower of Hanoi and London tasks. *Brain and Cognition, 56*(3): 286–292. https://doi.org/10.1016/j.bandc.2004.07.003

7 Original Manuscripts

7.1 Study 1: Validation of new tablet-based problem-solving tasks in primary school students

Schäfer, J., Reuter, T., Leuchter, M., & Karbach, J. (2024). Validation of new tablet-based problem-solving tasks in primary school students. *PLoS ONE*, *19*(8), e0309718. https://doi.org/10.1371/journal.pone.0309718

PLOS ONE

RESEARCH ARTICLE

Validation of new tablet-based problem-solving tasks in primary school students

Jonas Schäfer[1,2]*, **Timo Reuter**[2], **Miriam Leuchter**[2], **Julia Karbach**[1,3]

1 Department of Psychology, University of Kaiserslautern-Landau, Landau, Germany, **2** Institute for Child and Youth Education, University of Kaiserslautern-Landau, Landau, Germany, **3** Center for Research on Individual Development and Adaptive Education of Children at Risk (IDeA), Frankfurt, Germany

* jonas.schaefer@rptu.de

Citation: Schäfer J, Reuter T, Leuchter M, Karbach J (2024) Validation of new tablet-based problem-solving tasks in primary school students. PLoS ONE 19(8): e0309718. https://doi.org/10.1371/journal.pone.0309718

Editor: Adel Tekari, Centre of Biotechnology of Sfax: Centre de Biotechnologie de Sfax, TUNISIA

Received: March 12, 2024

Accepted: August 16, 2024

Published: August 29, 2024

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pone.0309718

Data Availability Statement: The data that support the findings of this study are openly available from the GitHub database in the repository GTT_Validation under the following URL: https://github.com/jonato-bit/GTT_Validation.git.

Abstract

Problem-solving is an important skill that is associated with reasoning abilities, action control and academic success. Nevertheless, empirical evidence on cognitive correlates of problem-solving performance in childhood is limited. Appropriate assessment tools are scarce and existing analog tasks require extensive coding. Thus, we developed and validated new tablet-based versions of existing analog tasks assessing technical problem-solving with gear construction tasks. To validate these tasks, 215 children (6–8 years) performed the problem-solving tasks in both modalities (analog, digital). To investigate whether performances in both modalities were correlated with other cognitive abilities, participants performed three additional tasks assessing language, reasoning and problem-solving. Structural equation modelling showed that performance was substantially correlated across modalities and also correlated with language, reasoning and another problem-solving task, showing the convergent validity of the digital tasks. We also found scalar measurement invariance across task modalities indicating that both task versions can be used interchangeably. We conclude that both versions (analog and digital) draw on similar cognitive resources and abilities. The analog tasks were thus successfully transferred to a digital platform. The new tasks offer the immense benefits of digital data collection, provide a valid measuring tool advancing problem-solving research in childhood and facilitate the application in the field, e.g., in the classroom.

Introduction

Problem-solving refers to the process of achieving a goal state that is different from an initial state by performing a series of cognitive or motor actions [1]. This process is typically characterized by the following interdependent phases: (a) understanding and mentally representing the problem, (b) developing plans and strategies to solve it, (c) practically implementing the plans and (d) evaluating (intermediate) outcomes [2]. Aside from minor variations, there is much agreement on these four phases of problem-solving across different domains [3–5]. Problem-solving skills develop significantly during the pre- and primary school years and benefit many important life outcomes, e.g., academic success and social skills [6]. Understanding,

Funding: This work was supported by the Ministry of Science, Further Education, Research and Culture, Rhineland-Palatinate, Germany. The funders awarded ML and JK (without grant number). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

structuring and intentionally solving a problem requires perceptual, motor and creative resources [7]. Exposure to problem-solving challenges promotes reasoning and learning processes in children [8] and is associated with systematic and critical thinking [9]. Therefore, problem-solving abilities are considered major educational learning outcomes [4]. Furthermore, problem-solving is closely related to higher-order cognitive functions, such as fluid intelligence [10] and executive functions [11].

Technical problem-solving is considered a prototypical subtype of general problem-solving [4], since it implements the four phases of problem-solving by performing manual actions (phase c) with observable effects (phase d) that are clearly different from the cognitively-based planning (phases a and b) [2, 12]. Therefore, technical problems have often been applied in recent empirical research on problem-solving in pre- and primary school children [13–15].

Since problem-solving relies on various higher-order cognitive processes [7, 16], the requirements for age-appropriate measures assessing problem-solving skills and strategies in children are high [17]. The design cognition framework focusses on the cognitive processes involved in technical designing and problem-solving and is commonly applied in pre- and primary school research [15]. Design cognition research identifies the problem-solving phases based on participants' think-aloud utterances during design task performance (e.g., "Design and build a bug box that does not allow frogs in but allows bugs in/out" [15]). Thus, conclusions on cognitive processes during technical problem-solving are commonly drawn from protocol-coding of children's actions and their own verbal reasoning on them [15].

To examine technical problem-solving skills in preschool children, previous studies applied gear turning tasks (*GTT*) with toy-like gears, propellers and a plastic pegboard [18]. In these tasks, children were to assemble gears and propellers according to specific instructions in terms of their turning direction and turning speed (e.g., assembling gears so that they would turn in the same direction). In contrast to protocol-based evaluation methods, this study design provided the advantage that correct target states were clearly defined and the phases of problem-solving could be observed by analyzing the way participants organized, moved, interconnected and turned the gears and propellers [18]. Moreover, it integrated relevant scientific knowledge with logical understanding and goal-directed behavior. However, the analog implementation of the GTTs required considerable time in terms of set-up, coding and data processing.

Compared to analog testing, digital assessment methods offer numerous advantages [19]. Performance data, such as reaction times, can be saved more precisely and reliably. Moreover, coding algorithms can process data efficiently, rendering the need for the very time-consuming analog coding obsolete. Only a terminal device is needed to perform the experiment, instead of large quantities of analog materials. Instructions can be presented more standardized and their timing is more comparable across subjects, while the tasks can be designed in a child-friendly way [20]. In current research, digital measurement is increasingly preferred over analog measurement. For instance, formerly paper-based tests are successively digitized in neuropsychological contexts to enable more efficient diagnostic assessments [21]. Similarly, problem-solving skills in adolescents are assessed by means of digital paradigms, such as *microworlds*, simulating real-life problem situations [22, 23], and classic cognitive tests such as the Corsi Block-Tapping task, the Stroop task and the Trail-Making test are routinely administered digitally across a wide range of ages [24].

However, newly digitized test instruments require thorough validation [24], which is usually achieved by demonstrating convergent validity of the digital test against the analog counterpart [25]. Convergent validity of digital multiple-choice tests is usually shown by significant medium-sized correlations with the paper-and-pencil version [26]. When measuring more complex digitized tests assessing higher-order cognitive skills, like the Trail-Making test [27], evidence for convergent validity is additionally provided by testing the strength of associations between related cognitive measures and both the digital and the analog versions of the test [28]. Moreover, the convergent validity is demonstrated via structural equation modelling assessing measurement invariance across latent factors of the analog and digital tests [29].

In this study, we developed digital, tablet-based versions of technical problem-solving tasks for children. We chose the GTTs that were applied in previous studies using analog test materials because they are appropriate for assessing five- to eight-year-old children's technical problem-solving [18]. Our aim was to validate the newly developed digital GTTs in six- to eight-year-old children. We therefore tested (1) whether performances on both modalities (analog and digital) were correlated with each other and with related measures (a language task, a reasoning task and another problem-solving task), (2) whether performance was measurement invariant across modalities, (3) whether performances on both modalities were predicted by related measures and (4) whether these predictive values differed between modalities. We expected substantial correlations between modalities, significant predictive values of related measures, that did not differ between modalities and that performances on both modalities (i.e., both task versions) were measurement invariant at the scalar level.

Materials and methods

Participants

A power-analysis (parameters: $r = .25$, $\alpha = .05$, $1\text{-}\beta = .90$) for correlational analyses resulted in a required sample size of $n = 164$. Power-analyses (parameters: $r = .25$, $\alpha = .05$, $1\text{-}\beta = .90$) for structural equation modeling analyses resulted in a required sample size of $n = 43$ for the highest degrees of freedom ($df = 7$) and $n = 54$ for the lowest degrees of freedom ($df = 5$). In sum, 215 children between six and eight years of age ($M = 7.18$ years, $SD = 0.78$; 89 female) participated in the study voluntarily and with written informed consent of their parents or caregivers. The recruitment period started on October 11, 2021 and ended on January 20, 2022. The study was approved by the local ethics committee (application #361). Participants were recruited in a town in southwestern Germany with a heterogeneous and diverse population [30]. The inclusion criterion was age (6–8 years; established prior to data analysis). We established no further exclusion criteria in order to minimize any selectivity of the sample.

Procedure

Participants completed ten tasks across two 30-minute sessions with a 60-minute break. During both sessions, the child worked individually with a trained experimenter at the lab. In session one, children completed a language task, a reasoning task and a problem-solving task (*stabilization task*) to assess the convergent validity of the newly developed digital problem-solving tasks. Children also completed three conceptual knowledge tasks that are not relevant for the present analyses. Session two served to validate the new digital version of the problem-solving gear turning tasks (GTTs) against the analog version. It therefore included both GTTs (*carousel* and *propeller*) in both modalities (analog, digital). Task modality (analog, digital) and task type (carousel, propeller) were counterbalanced across participants. Sessions were video recorded and tasks were instructed in German language.

Materials

All digital tasks were administered on a 10.1-inch sized tablet (Samsung Galaxy Note 10.1, Android version 5.1.1). The tablet program was developed in *Unity* (2020.3.17f1) and built via Android SDK (compression method LZ4). The digital GTTs were processed by single-point

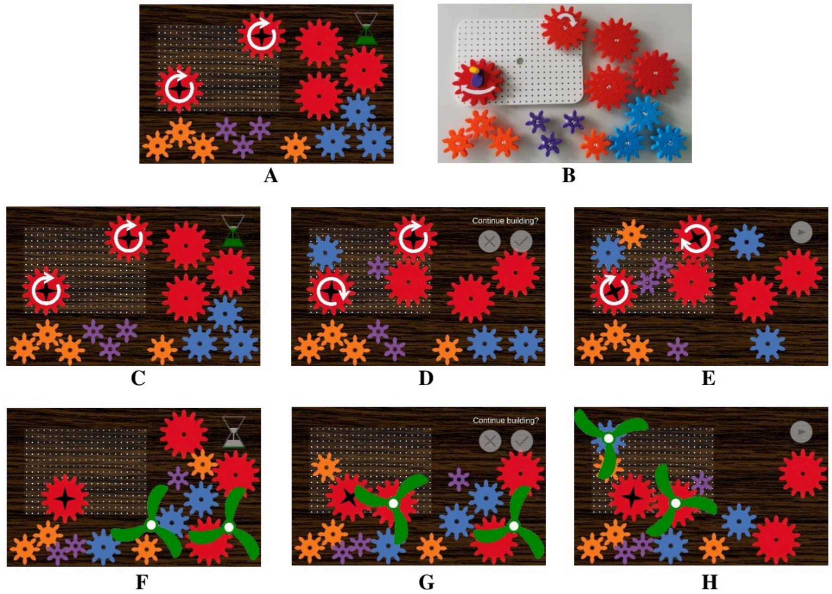

Fig 1. Experimental set-up of the GTTs. (A) Experimental set-up of the digital carousel task, (B) experimental set-up of the analog carousel task, (C) experimental phase 1 of the carousel task, (D) experimental phase 2 of the carousel task, (E) experimental phase 3 of the carousel task, (F) experimental phase 1 of the propeller task, (G) experimental phase 2 of the propeller task, (H) experimental phase 3 of the propeller task. GTT = Gear turning task.

https://doi.org/10.1371/journal.pone.0309718.g001

swipe control in a two-dimensional response space (Fig 1A). Each gear had an inner and an outer radius. The inner radius was defined as the distance between a gear's center and the closest point of the notches between two gear teeth. The outer radius extended to the outer end of the gear teeth. The gears and propellers were moved via drag-and-drop with the currently selected object slightly protruding in perspective. When a gear was dropped onto the board, its new position was aligned with a grid that exactly mimicked the analog pegboard (20x14 plug-in options). Since the inner radii of two gears on the board could not overlap, gears released at an invalid position snapped in at the position with the smallest possible distance to the release position that was not less than the sum of the two overlapping inner gear radii. Gears located on the board could be turned by circularly swiping in the area between the inner and the outer radius. Whenever the distance between two gears was less than the sum of their outer radii and greater than the sum of their inner radii, they were considered to be connected and, thus, drove each other when turning. Propellers were dragged to the center of a gear to attach them to it and consequently moved and turned uniformly with this gear. Children had sufficient time to familiarize with the tablet handling. The analog version of the GTTs, carousel and propeller, was administered with a pegboard and plastic gears (Fig 1B).

Tasks

Language: Receptive vocabulary (Wechsler Preschool and Primary Scale of Intelligence (WPPSI-IV)) [31]. Participants saw consecutive displays with four pictures on a tablet screen and had to point to the picture named by the experimenter. The task included two practice trials and up to 35 test trials. It was aborted after three consecutive errors. The dependent variable was the number of correctly solved items.

Reasoning: Matrix reasoning (WPPSI-IV [31]). Participants saw a 2x2 matrix on a tablet screen containing three pictures and one question mark. They were instructed to pick one out of four or five figurative response options to complete the matrix pattern. The task included three practice trials and up to 26 test trials. It was aborted after three consecutive errors. The dependent variable was the number of correctly solved items.

Problem-solving: Stabilization task (cf. [32]). Participants saw instable constructions of rectangular and triangular wooden blocks. Participants had up to three attempts to place another color-coded block stabilizing the construction. The task included one practice trial and eight test trials. The dependent variable was scored according to the number of required stabilizing attempts (items successfully solved at the first attempt: three points; one point less for each failed attempt; scoring range: 0–24).

Problem-solving: Carousel task [18]. Participants saw a rectangular board with a driving gear and a target gear (introduced as a carousel) on it (Fig 1C). Both were marked by a circular arrow indicating a clockwise turning direction. Participants had to make the target gear turn clockwise when turning the driving gear clockwise by using other gears of four different sizes to connect driving and target gear. After three minutes or when participants indicated that they had finished constructing, the first experimental phase was over. In the second experimental phase, the experimenter removed the pegboard (analog modality) or the tablet (digital modality) temporarily and showed a printed picture of the initial state of the board to the participants (i.e., the state shown in Fig 1C). Subsequently, participants were asked to repeat the task requirements. If they were not able to repeat them correctly, the experimenter repeated them. The accuracy of the participants' responses was coded video-based, but not analyzed in this study. Afterwards, participants saw their construction and decided whether they wanted to make further changes to it or to end the task if they felt that it was completed. If they continued, they had up to two minutes to finish the construction (third experimental phase; Fig 1E). The dependent variable was the solution quality of the final construction state (scoring system: no gear was connected to either of the fixed gears (driving and target gear): 0 points; one of the fixed gears was connected to at least one other gear: 1 point; both fixed gears were connected to at least one other gear: 2 points; the two fixed gears were connected, but did not turn in the same direction: 3 points; they turned in the same direction: 4 points).

Problem-solving: Propeller task [16]. Participants saw a rectangular board with an unmarked driving gear on it. In addition to different-sized gears, two propellers were provided (Fig 1F). Children had to attach them to gears in a way that one propeller would turn as fast as possible and the other one as slow as possible without touching each other when the driving gear was turned. Subsequently, the second and third experimental phase (shown in Fig 1G and 1H) proceeded in the same way as in the carousel task. The dependent variable was the solution quality of the final construction state. It was defined as the sum of two variables: *turning speed* (scoring system: at least one propeller was not attached to any gear: 0 points; both propellers were attached to the same gear type (i.e., gears of the same size): 1 point; both propellers were attached to different gear types, but not largest and smallest: 2 points; propellers were attached to the largest and smallest gear type: 3 points) and *contact* (scoring system: at least one propeller was not driven by the driving gear: 0 points; propellers touched when turning the driving gear: 1 point; they did not touch: 2 points).

Analyses

Performance on the GTTs was evaluated by two trained raters coding the solution quality by rating the same dependent measures that were collected and calculated by the tablet program in the digital version. The mean inter-rater reliability between the raters for data of 20 randomly selected participants was very good for each dependent variable (Cohen's Kappa = .82-.97). Given that the correlation between the manually coded data of the digital GTTs and the data measured by the tablet app was very substantial (carousel: $r = .87$, $p = .00$; propeller: $r = .94$, $p = .00$), we calculated the analyses based on the data measured by the tablet app.

Data of the analog GTTs of twelve participants was missing due to missing video recordings. Data of the analog propeller task of 21 participants was excluded because at least one propeller was not attached to the center but to the edge of a gear rendering the coding of turning speed impossible.

Statistical analyses were performed using R version 4.3.0 [33]. To analyze the manifest association between performances on both GTT modalities and other tasks, we used Spearman rank correlation analyses. As the main analysis, we estimated a model comprising two latent modality factors (analog and digital), each represented by performances on both GTTs (carousel and propeller task) of the respective modality. We fitted the covariance-based structural equation model with the R package lavaan [34] using only complete data sets. We evaluated goodness of fit based on the comparative fit index (*CFI*) [35] and the root mean square error of approximation (*RMSEA*) [36]. According to Hu and Bentler (1999), we considered CFI values >0.95 and RMSEA values <0.06 to indicate good model fit [37]. Additionally, we evaluated the standardized root mean square residual (*SRMR*) and the Tucker-Lewis index (*TLI*). According to Vermeent et al. (2022), we assessed longitudinal measurement invariance to control for configural (equal factor structure), metric (equal factor loadings) and scalar (equal intercepts) invariance across modalities using the likelihood ratio test [29]. Additionally, we included regression paths from the manifest indicators of language ability, reasoning ability and stabilization task performance to both modality factors in order to test their predictive value for GTT performance (convergent validity). We then imposed equality constraints on the regression paths in order to check for differences in this predictive value between modalities.

Results

Descriptives and correlation coefficients for the dependent variables are presented in Table 1.

Correlations between GTT modalities and other tasks

On the manifest level, we found significant associations between modalities in both tasks (carousel: $r = .33$, $p = .00$; propeller: $r = .50$, $p = .00$). Performance across all four GTTs correlated significantly with performance on the language task (carousel analog: $r = .31$; carousel digital: r

Table 1. Means and standard deviations of the GTTs (carousel, propeller) and the validation tasks (reasoning, language and problem-solving) and correlations between these measures.

Variable	*M*	*SD*	1	2	3	4	5	6
1. Language task	27.02	4.44						
2. Reasoning task	16.20	4.47	.41**					
3. Problem-solving task	12.22	3.20	.37**	.36**				
4. GTT: Carousel analog	3.35	1.05	.31**	.31**	.32**			
5. GTT: Carousel digital	3.06	1.18	.34**	.43**	.28**	.33**		
6. GTT: Propeller analog	3.15	1.57	.20**	.24**	.42**	.35**	.28**	
7. GTT: Propeller digital	2.43	1.66	.38**	.34**	.36**	.24**	.46**	.50**

Note: GTT = Gear turning task;

*$p < .05$,

**$p < .01$.

https://doi.org/10.1371/journal.pone.0309718.t001

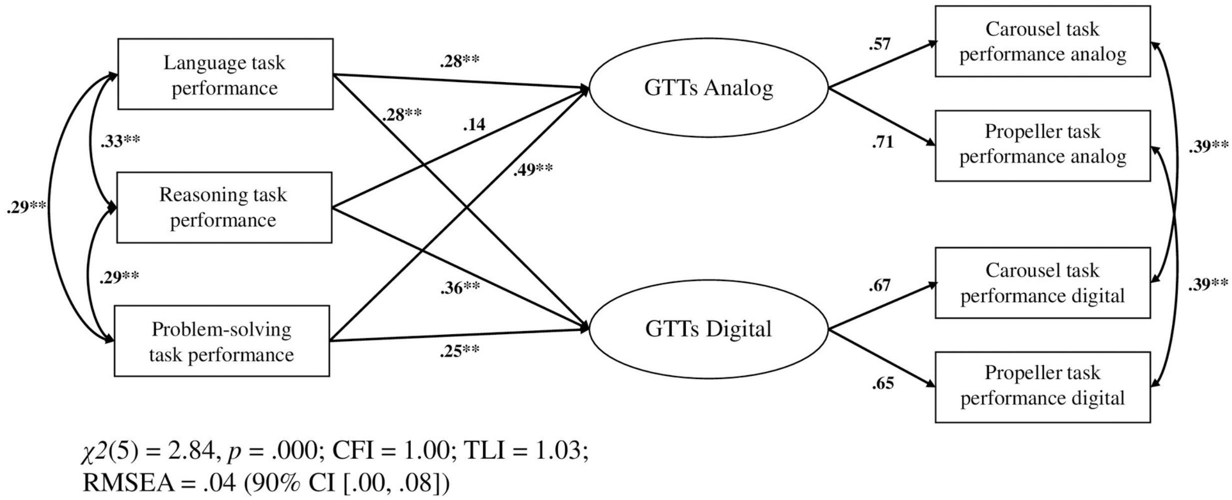

Fig 2. Model describing the relationships of analog and digital GTT performances and related measures (language, reasoning and stabilization problem-solving). The squares represent manifest variables and the circles represent latent variables. Single-headed arrows from manifest variables to latent variables indicate regression paths, single-headed arrows from latent variables to manifest variables indicate factor loadings, double-headed arrows indicate correlations. The model was estimated with weighted least square mean and variance adjusted (WLSMV) for ordinal data. GTT = Gear turning task; $^{*}p < .05$, $^{**}p < .01$.

https://doi.org/10.1371/journal.pone.0309718.g002

= .34; propeller analog: $r = .20$; propeller digital: $r = .38$; all $p < .01$). Performance on all GTTs was also significantly correlated with performance on the reasoning task (carousel analog: $r = .31$; carousel digital: $r = .43$; propeller analog: $r = .24$; propeller digital: $r = .34$; all $p < .01$). Moreover, performance on the GTTs was significantly correlated with performance on the stabilization task (carousel analog: $r = .32$; carousel digital: $r = .28$; propeller analog: $r = .42$; propeller digital: $r = .36$; all $p < .01$).

Measurement invariance across GTT modalities

The latent factor model provided an excellent fit to the data (see Fig 2). Likelihood ratio tests revealed configural, metric and scalar measurement invariance across the modality factors (analog and digital; see Table 2).

Predictive value of performance on related tasks for GTT performance (convergent validity)

Regression paths in the model (see Fig 2) were significant for language and stabilization performance across modalities. Reasoning was a significant predictor of digital GTT performance,

Table 2. Model parameters of measurement invariance tests.

Level	$\chi2(df)$	p	Robust CFI	Robust TLI	Robust RMSEA [90% CI]	SRMR	$\Delta\chi2$	$p(\Delta\chi2)$
Configural	2.839 (5)	.000	1.000	1.027	.038 [.000-.082]	.027		
Metric	3.449 (6)	.000	1.000	1.027	.034 [.000-.080]	.032	0.610	.326
Scalar	3.890 (7)	.000	1.000	1.028	.033 [.000-.075]	.032	0.441	.284

Note: Models were estimated with weighted least square mean and variance adjusted (WLSMV) for ordinal data. CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; CI = Confidence interval; SRMR = Standardized Root Mean Square Residual.

https://doi.org/10.1371/journal.pone.0309718.t002

Table 3. Comparison of regression paths between task modalities (analog, digital).

Model	$\chi2(df)$	p	Robust CFI	Robust TLI	Robust RMSEA [90% CI]	SRMR	$\Delta\chi2$	$p(\Delta\chi2)$
Unconstrained	2.839 (5)	.000	1.000	1.027	.038 [.000-.082]	.027		
Language	2.972 (6)	.000	1.000	1.032	.033 [.000-.074]	.028	0.133	.480
Reasoning	5.365 (6)	.000	1.000	1.007	.055 [.012-.096]	.039	2.526	.047*
Stabilization task	3.990 (6)	.000	1.000	1.021	.041 [.000-.085]	.032	1.151	.190

Note: All three regression paths were constrained separately to be equal across task modalities. Models were estimated with weighted least square mean and variance adjusted (WLSMV) for ordinal data. CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; CI = Confidence interval; SRMR = Standardized Root Mean Square Residual.

*$p < .05$,

**$p < .01$.

https://doi.org/10.1371/journal.pone.0309718.t003

but not of analog GTT performance. Imposing equality constraints on the regression paths between modalities did not significantly change the model fit for language and stabilization performance, but for reasoning (see Table 3). Thus, language and stabilization performance contributed equally strong to GTT performances across modalities, whereas reasoning contributed more strongly to digital than to analog GTT performance.

Discussion

Our aim was to validate new, digital versions of gear-based problem-solving tasks against the traditional analog tasks. Six- to eight-year-old children performed both the analog and the digital version of two GTTs as well as a number of related validation measures. First, we compared performance between both modalities (digital, analog). As expected, we found significant correlations between task performances across modalities. These results are consistent with findings from earlier studies [e.g., 24, 27, 38] that showed medium-sized correlations between digitized and analog versions of cognitive tests, indicating that both versions are comparable. Second, we found measurement invariance across the digital and the analog GTT versions on the configural, metric and scalar level. This indicates that both task versions had an equal factor structure, equal factor loadings and an equal scale level (i.e., equal intercepts), which also provides evidence for convergent validity. Third, we found the expected predictive value of language, reasoning and a problem-solving task performance for GTT performances on both modalities, providing further evidence for convergent validity of the tasks. Fourth, this predictive value was comparable across task modalities for language and stabilization performance, but larger in the digital modality for reasoning. The association between GTT performance and language ability indicates that children substantially relied on verbal processes to solve the GTT problems. In fact, this is in line with the assumption that children use subvocal verbal self-instruction to solve complex tasks [39–42]. The substantial association between GTT performance and stabilization task performance indicates that the new tasks are construct valid and might be representative for multiple problem-solving domains [43]. However, reasoning contributed more strongly to the digital than to the analog GTT performance. Since previous studies have shown that reasoning is an essential component of problem-solving [22, 44], this positive association confirmed the construct validity of the new digital measurement. The fact that reasoning was not predictive for the analog GTT performance might have been caused by a modality effect, since the reasoning test was also assessed digitally on a tablet.

The many advantages of digital testing [45] include that it can provide more accurate and detailed data and record multiple measurements without the need for trained personnel to perform video-based data coding (also eliminating the need to film hours of testing and get parental consent for video recordings, increasing the efficiency of testing). The reliability of the digital data processing is clearly documented by very high correlations between human-coded data and data measured by the app. Furthermore, digital assessment methods are much easier to implement outside of the lab (e.g., in schools or daycare centers) and reduce test administrator influence on the assessment results [46]. Although the benefits of digital problem-solving assessment have been known for a long time [19], so far only very few digital assessment tools for children are available.

Regarding the framework of problem-solving, our study focused on solution quality rather than on the individual problem-solving phases [2]. Many previous studies lacked valid quantitative measures of solution quality, which may be of particular importance for the design of intervention studies and educational support programs. The digital GTTs automatically track both the solution quality as well as data providing information on the process of problem-solving, including the number and timing of gear turnings, gear displacements and processing time. This process data provides opportunities to closer investigate the cognitive processes (e.g., reasoning processes) involved in the different problem-solving phases. Therefore, the new tasks provide a significant improvement in the measurement of technical problem-solving strategies beyond qualitative approaches [15, 18] and extend current quantitative accuracy-based research on children's problem-solving [22, 23]. The digital GTTs could furthermore be adapted for the assessment of older children and adolescents by adjusting the difficulty level of the tasks, for example, by increasing distances between fixed gears or reducing the number of available gears. A limitation of our study is the lack of more specific information about the sample, such as the socio-economic status of the participants' caregivers.

Conclusion

We conclude that the new digital GTTs are a valid adaptation of the traditional analog tasks for assessing problem-solving in six- to eight-year-old children. Given the substantial advantages of digital assessment instruments, the development of this tablet app contributes to research on cognitive development in problem-solving skills and strategies in early and middle childhood. The new tablet-based tasks also provide opportunities to investigate the contribution of higher-order cognitive abilities, such as executive functions, to problem-solving abilities in childhood [47]. Further studies are needed to collect representative normative data for the digital GTTs considering, for example, the effects of age, grade and gender.

Acknowledgments

We especially want to thank Steffen Pohl and Rebekka Morath for their valuable assistance in the preparation of the study. Furthermore, we would like to thank the student research assistants for their support in data collection.

Author Contributions

Conceptualization: Jonas Schäfer, Timo Reuter, Miriam Leuchter, Julia Karbach.

Data curation: Jonas Schäfer.

Formal analysis: Jonas Schäfer.

Funding acquisition: Miriam Leuchter, Julia Karbach.

Investigation: Jonas Schäfer, Timo Reuter.

Methodology: Jonas Schäfer, Julia Karbach.

Project administration: Jonas Schäfer.

Resources: Miriam Leuchter, Julia Karbach.

Software: Jonas Schäfer.

Supervision: Julia Karbach.

Validation: Jonas Schäfer.

Visualization: Jonas Schäfer.

Writing – original draft: Jonas Schäfer.

Writing – review & editing: Jonas Schäfer, Timo Reuter, Miriam Leuchter, Julia Karbach.

References

1. OECD. The nature of problem solving: Using research to inspire 21st century learning. Csapó B, Funke J, editors. Paris: OECD Publishing; 2017.
2. Polya G. Patterns of plausible inference. Princeton University Press; 1954.
3. Dostál J. Theory of problem solving. Procedia Soc Behav Sci. 2015; 174: 2798–2805. https://doi.org/10.1016/j.sbspro.2015.01.970
4. Jonassen DH. Toward a design theory of problem solving. Educ Technol Res Dev. 2000; 48: 63–85. https://doi.org/10.1007/BF02300500
5. Soebagyo J, Habibie H, Gunawan I, editors. Polya's four phases exploration in solving linear program story questions based on student beliefs. Proceedings of the 1st Annual International Conference on Natural and Social Science Education (ICNSSE 2020); 2021;547: 260–267.
6. Diamond LL. Problem solving in the early years. Interv Sch Clin. 2018; 53(4): 220–223. https://doi.org/10.1177/1053451217712957
7. Wang Y, Chiew V. On the cognitive process of human problem solving. Cogn Syst Res. 2010; 11(1): 81–92. https://doi.org/10.1016/j.cogsys.2008.08.003
8. George J, Greenfield DB. Examination of a structured problem-solving flexibility task for assessing approaches to learning in young children: Relation to teacher ratings and children's achievement. J Appl Dev Psychol. 2005; 26(1): 69–84. https://doi.org/10.1016/j.appdev.2004.10.006
9. Rahman MM. 21st Century skill "problem solving": Defining the concept. Asian J Interdiscip Res. 2019;64–74. https://doi.org/10.34256/ajir1917
10. Zook NA, Davalos DB, DeLosh EL, Davis HP. Working memory, inhibition, and fluid intelligence as predictors of performance on Tower of Hanoi and London tasks. Brain Cogn. 2004; 56(3): 286–292. https://doi.org/10.1016/j.bandc.2004.07.003 PMID: 15522766
11. Karbach J, Unger K. Executive control training from middle childhood to adolescence. Front Psychol. 2014; 5: 390. https://doi.org/10.3389/fpsyg.2014.00390 PMID: 24847294
12. Keen R. The development of problem solving in young children: A critical cognitive skill. Annu Rev Psychol. 2011; 62: 1–21. https://doi.org/10.1146/annurev.psych.031809.130730 PMID: 20822435
13. Gold ZS, Elicker J, Kellerman AM, Christ S, Mishra AA, Howe N. Engineering play, mathematics, and spatial skills in children with and without disabilities. Early Educ Dev. 2021; 32(1): 49–65. https://doi.org/10.1080/10409289.2019.1709382
14. Shechter T, Eden S, Spektor-Levy O. Preschoolers' nascent engineering thinking during a construction task. J Cogn Educ Psychol. 2021; 20(2): 83–111. https://doi.org/10.1891/JCEP-D-20-00010
15. Strimel GJ, Bartholomew SR, Kim E, Zhang L. An investigation of engineering design cognition and achievement in primary school. J STEM Educ Res. 2018; 1: 173–201. https://doi.org/10.1007/s41979-018-0008-0
16. Schäfer J, Reuter T, Karbach J, Leuchter M. Domain-specific knowledge and domain-general abilities in children's science problem-solving. Br J Educ Psychol. 2024; 94(2): 346–366. https://doi.org/10.1111/bjep.12649 PMID: 38061789
17. Greiff S, Fischer A. Measuring complex problem solving: An educational application of psychological theories. J Educ Res online. 2013; 5(1): 38–58. https://doi.org/10.25656/01:8019
18. Reuter T, Leuchter M. Examining kindergarten children's testing and optimising in the context of a gear engineering task. European Journal of STEM Education. 2022; 7(1): 04. https://doi.org/10.20897/ejsteme/11827

19. Baker EL O'Neil HF. Measuring problem solving in computer environments: current and future states. Comput Human Behav. 2002; 18(6): 609–622. https://doi.org/10.1016/S0747-5632(02)00019-5
20. Johann VE, Karbach J. Validation of new online game-based executive function tasks for children. J Exp Child Psychol. 2018; 176: 150–161. https://doi.org/10.1016/j.jecp.2018.07.009 PMID: 30119937
21. Latendorf A, Runde LM, Salminen T, Steinert A. Digitization of neuropsychological diagnostics: A pilot study to compare three paper-based and digitized cognitive assessments. Aging Clin Exp Res. 2021; 33: 1585–1597. https://doi.org/10.1007/s40520-020-01668-z PMID: 33001402
22. Greiff S, Krkovic K, Hautamäki J. The prediction of problem-solving assessed via microworlds. Eur J Psychol Assess. 2016; 32(4): 298–306. https://doi.org/10.1027/1015-5759/a000263
23. Krieger F, Stadler M, Bühner M, Fischer F, Greiff S. Assessing complex problem-solving skills in under 20 minutes. Psychological Test Adaptation and Development. 2021; 2(1): 80–92. https://doi.org/10.1027/2698-1866/a000009
24. Björngrim S, van den Hurk W, Betancort M, Machado A, Lindau M. Comparing traditional and digitized cognitive tests used in standard clinical evaluation—A study of the digital application Minnemera. Front Psychol. 2019; 10: 2327. https://doi.org/10.3389/fpsyg.2019.02327 PMID: 31681117
25. Kiili K, Ketamo H. Evaluating cognitive and affective outcomes of a digital game-based math test. IEEE Trans Learn Technol. 2018; 11(2): 255–263. https://doi.org/10.1109/TLT.2017.2687458
26. Cattaneo G, Pachón-García C, Roca A, Alviarez-Schulze V, Opisso E, García-Molina A, et al. "Guttmann Cognitest"®, preliminary validation of a digital solution to test cognitive performance. Front Aging Neurosci. 2022; 14: 987891. https://doi.org/10.3389/fnagi.2022.987891 PMID: 36408102
27. Baykara E, Kuhn C, Linz N, Tröger J, Karbach J. Validation of a digital, tablet-based version of the Trail Making test in the Δelta platform. Eur J Neurosci. 2022; 55(2): 461–467. https://doi.org/10.1111/ejn.15541 PMID: 34811827
28. Hittner JB, May K, Silver NC. A Monte Carlo evaluation of tests for comparing dependent correlations. J Gen Psychol. 2003; 130(2): 149–168. https://doi.org/10.1080/00221300309601282 PMID: 12773018
29. Vermeent S, Spaltman M, van Elswijk G, Miller JB, Schmand B. Philips IntelliSpace Cognition digital test battery: Equivalence and measurement invariance compared to traditional analog test versions. Clin Neuropsychol. 2022; 36(8): 2278–2299. https://doi.org/10.1080/13854046.2021.1974565 PMID: 34528868
30. Ackel-Eisnach K. Landau in Zahlen—Analyse bevölkerungsstatistischer Indikatoren zur gesellschaftlichen Vielfalt. In: Pusch B, Spieker S, Horne C. (eds) Heterogenität und Diversität in Städten mittlerer Größe. Wiesbaden: Springer; 2023.
31. Petermann F, Daseking M. Wechsler preschool and primary scale of intelligence—Fourth edition. Göttingen: Hogrefe; 2018.
32. Weber AM, Leuchter M. Measuring preschool children's knowledge of the principle of static equilibrium in the context of building blocks: Validation of a test instrument. Br J Educ Psychol. 2020; 90: 50–74. https://doi.org/10.1111/bjep.12304 PMID: 31292951
33. R Core Team. R: A Language and Environment for Statistical Computing. Version 4.3.0 [software]. R Foundation for Statistical Computing. 2023 Apr 21 [cited 2024 Mar 10] https://www.R-project.org/
34. Rosseel Y. lavaan: An R package for structural equation modeling. J Stat Softw. 2012; 48(2): 1–36. https://doi.org/10.18637/jss.v048.i02
35. Bentler PM. Comparative fit indexes in structural models. Psychol Bull. 1990; 107(2): 238–246. https://doi.org/10.1037/0033-2909.107.2.238 PMID: 2320703
36. Browne MW, Cudeck R. Alternative ways of assessing model fit. Sociological Methods and Research. 1992; 21(2): 230–258. https://doi.org/10.1177/0049124192021002005
37. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equ Modeling. 1999; 6(1): 1–55. https://doi.org/10.1080/10705519909540118
38. Fellows RP, Dahmen J, Cook D, & Schmitter-Edgecombe M. Multicomponent analysis of a digital Trail Making Test. Clin Neuropsychol. 2017; 31(1): 154–167. https://doi.org/10.1080/13854046.2016.1238510 PMID: 27690752
39. Karbach J, Kray J. Developmental changes in switching between mental task sets: The influence of verbal labeling in childhood. J Cogn Dev. 2007; 8(2): 205–236. https://doi.org/10.1080/15248370701202430
40. Kray J, Eber J, Karbach J. Verbal self-instructions in task switching: A compensatory tool for action-control deficits in childhood and old age? Dev Sci. 2008; 11(2): 223–236. https://doi.org/10.1111/j.1467-7687.2008.00673.x PMID: 18333979

41. Gunzenhauser C, Karbach J, Saalbach H. Function of verbal strategies in monolingual vs. bilingual students' planning performance: An experimental approach. Cogn Dev. 2019; 50: 1–12. https://doi.org/10.1016/j.cogdev.2019.01.003

42. Enke S, Gunzenhauser C, Hepach R, Karbach J, Saalbach H. Differences in cognitive processing? The role of verbal processes and mental effort in bilingual and monolingual children's planning performance. J Exp Child Psychol. 2022; 213: 105255. https://doi.org/10.1016/j.jecp.2021.105255 PMID: 34388641

43. Greiff S, Wüstenberg S, Csapó B, Demetriou A, Hautamäki J, Graesser AC, et al. Domain-general problem solving skills and education in the 21st century. Educ Res Rev. 2014; 13: 74–83. https://doi.org/10.1016/j.edurev.2014.10.002

44. Kim M, Pegg J. Case analysis of children's reasoning in problem-solving process. Int J Sci Educ. 2019; 41(6): 739–758. https://doi.org/10.1080/09500693.2019.1579391

45. Germine L, Reinecke K, Chaytor NS. Digital neuropsychology: Challenges and opportunities at the intersection of science and software. Clin Neuropsychol. 2019; 33(2): 271–286. https://doi.org/10.1080/13854046.2018.1535662 PMID: 30614374

46. Overton M, Pihlsgård M, Elmståhl S. Test administrator effects on cognitive performance in a longitudinal study of ageing. Cogent Psychol. 2016; 3(1): 1260237. https://doi.org/10.1080/23311908.2016.1260237

47. Schäfer J, Reuter T, Leuchter M, Karbach J. Executive functions and problem-solving—the contribution of inhibition, working memory, and cognitive flexibility to science problem-solving performance in elementary school students. J Exp Child Psychol. 2024; 244: 105962. https://doi.org/10.1016/j.jecp.2024.105962 PMID: 38810499

7.2 Study 2: Domain-specific knowledge and domain-general abilities in children's science problem-solving

Schäfer, J., Reuter, T., Karbach, J., & Leuchter, M. (2024). Domain-specific knowledge and domain-general abilities in children's science problem-solving. *British Journal of Educational Psychology, 94*(2), 346–366. https://doi.org/10.1111/bjep.12649

Received: 27 July 2023 | Accepted: 23 November 2023

DOI: 10.1111/bjep.12649

ARTICLE

the british psychological society
promoting excellence in psychology

Domain-specific knowledge and domain-general abilities in children's science problem-solving

Jonas Schäfer[1,2] | Timo Reuter[1] | Julia Karbach[2,3] | Miriam Leuchter[1]

[1]Institute for Children and Youth Education, University of Kaiserslautern-Landau, Landau, Germany

[2]Department of Psychology, University of Kaiserslautern-Landau, Landau, Germany

[3]Center for Research on Individual Development and Adaptive Education of Children at Risk (IDeA), Frankfurt, Germany

Correspondence
Jonas Schäfer, August-Croissant-Straße 5, 76829 Landau, Germany.
Email: jonas.schaefer@rptu.de

Funding information
Ministeriums für Wissenschaft, Weiterbildung und Kultur, Rheinland-Pfalz

Abstract

Background: Problem-solving in early and middle childhood is of high relevance for cognitive developmental research and educational support. Previous research on science problem-solving has focussed on the process and strategies of children handling challenging tasks, but less on providing insights into the cognitive network that enables science problem-solving.

Aims: In this study, we aimed to investigate whether performance in science problem-solving is mainly determined by domain-specific rule knowledge, by domain-general cognitive abilities or both.

Methods: In our study, 215 6- to 8-year-old children completed a set of three domain-specific rule knowledge tasks and three corresponding problem-solving tasks that were content-coherent, as well as a vocabulary task, and a reasoning task.

Results: Correlational and regression analyses revealed a negligible impact of domain-specific rule knowledge on corresponding problem-solving tasks. In contrast, the associations between problem-solving performance in different domains and the associations between problem-solving performance and domain-general abilities (vocabulary and reasoning) were comparably strong.

Conclusions: The findings suggest that science problem-solving in primary school children primarily relies on domain-general cognitive abilities. Implications of these findings are discussed with regard to cognitive theories and early science education.

KEYWORDS

domain-general cognition, domain-specific knowledge, primary school education, problem-solving, rule knowledge

INTRODUCTION

Problem-solving abilities in children are assumed to foster subsequent personal and professional success (Diamond, 2018). Thus, problem-solving is a relevant research topic regarding childhood development and educational support (Keen, 2011). Problem-solving describes the process of reaching a desired state (i.e., a goal) that is different from an initial state (Funke et al., 2018). This process requires cognitive effort on the part of the problem solver because the gap between the initial state and the desired state arises from a lack of ability or routine knowledge (Funke et al., 2018; Mehadi Rahman, 2019). The problem-solving process includes interdependent steps, such as understanding the problem, testing partial solutions and continuously matching the current state with the goal state (Adams & Atman, 1999; Funke et al., 2018; Lee & Johnson-Laird, 2013; Strimel et al., 2018).

A core aspect of problem-solving is rule knowledge that highlights logical problem aspects (Luo & Niki, 2003), which can be related to domain-specific content. Some researchers consider domain-specific content knowledge as the main factor contributing to successful problem-solving (cf. the seminal work by Chi et al., 1982; Perkins & Salomon, 1989), while others consider domain-general aspects, such as fluid and crystallized intelligence, crucial for problem-solving (Greiff et al., 2014; Mehadi Rahman, 2019).

Our study aims to analyse the role of domain-specific rule knowledge and the role of domain-general cognitive abilities in 6- to 8-year-old children's problem-solving. For this purpose, we tested domain-specific rule knowledge in two content domains: gear-turning mechanisms and stability of building block constructions. In addition, we measured vocabulary and reasoning as indicators of intelligence and investigated the effects of age and gender on problem-solving performance.

The role of rule knowledge in problem-solving

Rule knowledge represents the understanding of relationships underlying logical problems, whereas rule application represents the use of this knowledge to solve a problem (Bühner et al., 2008; Fischer et al., 2011; Greiff et al., 2015; Jonassen, 2000; Mustakim et al., 2020; Simmons, 1992). Rule knowledge can be considered as domain-independent (Greiff et al., 2015) and mentally represented as a logical if-then-proposition (Arló-Costa & Levi, 1996). However, domain-specific rule knowledge can primarily be applied to a problem in this specific domain, for example in the gear domain. For instance, the domain-specific rule knowledge of turning direction can be applied to make gears turn in particular directions. Effectively applied domain-specific rule knowledge can be an integral part of successful problem-solving (Charlesworth & Leali, 2012; Milbourne & Wiebe, 2018).

Domain-specific rule knowledge can be stored in memory explicitly or implicitly (Reber & Kotovsky, 1997). A problem solver with explicit domain-specific rule knowledge can explain the rules, for example the turning direction of interconnected gears. A problem solver with implicit domain-specific rule knowledge knows about the turning direction of gears but is not able to explain it. Both explicit and implicit rule knowledge may be applied in problem-solving (Hamilton et al., 2007). Problem-solving without explicit knowledge is possible when using trial-and-error as a heuristic (Funke et al., 2018). Trial-and-error can be considered the lowest level of prior planning (Stern & Hertel, 2022), but can result in successful problem-solving by chance, lead to strategy formation, and generate implicit knowledge (Reber & Kotovsky, 1997; Tönnsen, 2021).

Domain-specific knowledge and domain-general cognitive abilities in problem-solving

Domain-specific knowledge and domain-general cognitive abilities in academic performance, including problem-solving, have been a much-investigated field for decades (Greene et al., 2018). However, findings on the interaction of domain-specific knowledge and domain-general cognitive abilities in problem-solving are inconsistent (Chi et al., 1981; Chiappe & MacDonald, 2005; Greiff et al., 2014).

Domain-specific content knowledge in problem-solving

Studies using the expert-novice paradigm indicated that more domain-specific content knowledge (expertise) correlates with enhanced problem-solving abilities in the respective domain (e.g., Chi et al., 1981; Perkins & Salomon, 1989). Consequently, students' context familiarity of a problem may improve their problem-solving (Bibi et al., 2018). Content knowledge allows experts to pay attention to the deep structure of a problem, whereas novices are typically oriented towards surface features (Chi, 2006; Mayer, 1997). Moreover, a certain amount of domain-specific content knowledge may be necessary for an effective application of domain-general problem-solving strategies (Alexander & Judy, 1988), such as problem decomposition and solution search (Reif & Heller, 1982). A recent study with physics students indicated that problem-solving performance increased as a function of domain-specific content knowledge (Milbourne & Wiebe, 2018). Yet, findings from Sabella and Redish (2007) suggested that performance in problem-solving not only depends on the availability of domain-specific content knowledge but more on how and when it is activated. Thus, domain-specific content knowledge does not completely explain problem-solving performance (e.g., Schoenfeld, 1987), suggesting that other cognitive processes are involved.

Domain-general cognitive abilities in problem-solving

A number of previous studies indicated that higher-order cognition, such as fluid and crystallized intelligence, contributes to problem-solving performance (Funke, 2014; Greiff et al., 2014; Mehadi Rahman, 2019; Zook et al., 2004). Fluid intelligence might, for instance, support problem-solving performance when complex problems have to be segmented or inductive or deductive reasoning is required (Chiappe & MacDonald, 2005; O'Brien et al., 2023). Crystallized intelligence is considered an important factor in children's problem-solving, as it allows the retrieval of previously acquired information and experiences, but also as a prerequisite for the understanding of verbal problem statements (Stephan et al., 2022). Specifically, verbal abilities support action planning, which is a central part of problem-solving (Gunzenhauser et al., 2019).

The interrelation of domain-specific content knowledge and domain-general cognitive abilities in children

There are only a few studies focusing on the relationship between domain-specific content knowledge and domain-general cognitive abilities in children's problem-solving. Findings from English (1992) indicated that the amount of domain-specific content knowledge plays a crucial role in 4- to 9-year-olds for the application of meaningful solution strategies to combinatorial problems. Gilmore et al. (2018) suggested that domain-specific content knowledge regarding the structure of the number system best predicted arithmetic problem-solving, whereas domain-general cognitive abilities were weaker predictors in 8- to 10-year-old children. In contrast, Geary et al. (2017) found in a longitudinal study from grade one to grade eight that domain-general cognitive abilities were more important than domain-specific content knowledge for mathematics problem-solving in early grades, but domain-general abilities and domain-specific content knowledge were equally important in later grades.

Besides the mathematics domain, children's problem-solving has been studied in an early engineering framework (for an overview, Gold & Elicker, 2020; Lucas & Hanson, 2016; Strimel et al., 2018). A study by Spektor-Levy and Shechter (2022) has shown that engineering practices and problem-solving performance of 5- and 6-year-old preschoolers improved by simply providing them with construction materials for free play at their daycare centre. However, few studies on children's engineering examined the role of domain-specific content knowledge for problem-solving. Reuter and Leuchter (2022a) investigated problem-solving with 4- to 7-year-olds in the domain of gears. The task involved connecting two gears fixed on a board so that they turn in the same direction. Based on the rule that adjacent gears turn in opposite directions, the children had to use an odd number of additional gears to successfully solve the problem (Reuter & Leuchter, 2022a). The results revealed no significant correlations between domain-specific rule knowledge (i.e., turning direction of adjacent gears) and quality of the problem solution (Reuter & Leuchter, 2022a).

Young children are likely to lack comprehensive content knowledge in various domains. However, they may have phenomenological primitives (diSessa, 1988) which they may have gained during playful experience (Weisberg et al., 2013). For example, the phenomenon of gears' turning direction and turning speed may have been achieved while playing with gears and the phenomenon of stability through block play. This experience may result in rule knowledge.

Children's domain-specific rule knowledge of gears' turning direction and turning speed

In the domain of gear-turning mechanisms, children may develop rule knowledge through playful experiences. The turning direction of adjacent connected gears is opposite. Thus, a target gear turns in the same direction as a driving gear when connected by an odd number of intermediate gears (Lehrer & Schauble, 1998). The turning speed of a gear around its own axis is size-dependent. Smaller gears turn faster around their own axis compared to larger gears (Reuter & Leuchter, 2021). Several studies have shown that children's domain-specific rule knowledge regarding gears' turning direction and turning speed develops during early school years and can be fostered through playful interventions (Lehrer & Schauble, 1998; Reuter & Leuchter, 2021, 2022a). In the study by Reuter and Leuchter (2021), most 7- and 8-year-olds had the correct domain-specific rule knowledge, whereas the majority of 5- and 6-year-olds did not. With respect to gears' turning speed, most children between 5 and 10 years of age had insufficient domain-specific rule knowledge (Reuter & Leuchter, 2021). However, studies by Reuter and Leuchter (2021, 2022b) indicated that 5- and 6-year-olds might acquire appropriate domain-specific rule knowledge about gears' turning direction and speed in short play-based interventions.

Children's domain-specific rule knowledge of building block constructions' stability

The stability of building blocks depends on the centre-of-mass principle, stating that a resting object is unstable when it is not supported (Weber & Leuchter, 2020). The geometric centre aligns with the centre-of-mass in symmetrical but not in asymmetrical constructions. Research indicated that asymmetrical building block constructions are more difficult to rate correctly than symmetrical building block constructions for 3- to 6-year-old children (Krist, 2010), indicating a lack of explicit stability knowledge. Additionally, Krist et al. (2005) found that 4- and 5-year-olds struggle to actively balance symmetrical building blocks on a beam scale, suggesting a lack of appropriate implicit knowledge. However, the development of stability knowledge between 4 and 8 years (Krist et al., 2005) can be fostered through interventions that address children's strategies to stabilize building blocks and their reasoning about it (Weber et al., 2020). These studies indicated that domain-specific rule knowledge can be learned implicitly and is accessible implicitly and explicitly.

Taken together, both the turning direction and speed of gears and the stability of building block constructions seem to be appropriate content domains to investigate the relationship between domain-specific rule knowledge and domain-general abilities for 6-to-8-year-olds' problem-solving. Thus, in the present study, we assessed problem-solving with tasks focusing on gears' turning direction and turning speed and building block constructions' stability.

The role of age and gender in problem-solving abilities

Problem-solving abilities significantly improve across childhood (Keen, 2011). Thornton (2009) attributed this learning development to children's active engagement with problem situations. Research suggested that problem-solving in children develops from trial-and-error procedures towards a more strategic and planful process with increasing age (Injoque-Ricle et al., 2014). However, age-related changes in problem-solving do not develop automatically but are driven by children's experiences in tool use, object manipulation and understanding of causal relations (Keen, 2011). Regarding general science understanding, a developmental boost occurs during early primary school age (Fitzgerald & Smith, 2016). Moreover, a previous study has shown that a developmental shift in the conceptual understanding of gears might occur between the ages of 6 and 7 years (Reuter & Leuchter, 2021). This was indicated by the fact that 7-year-olds were more likely than 6-year-olds to understand the content underlying gear problems (Reuter & Leuchter, 2021). While problem-solving strategies rely strongly on crystallized intelligence in older adults and on fluid intelligence in younger adults (Chen et al., 2017), less is known about early childhood.

Evidence for gender effects in problem-solving is heterogeneous, although there is some evidence that problem-solving performance varies between genders according to the problem domain (Walker et al., 2002; Zhu, 2007). Research suggested that female preschoolers are more competent in applying social problem-solving strategies than males (Walker et al., 2002). In contrast, male sixth graders were more strategic and successful in mathematical problem-solving (Zhu, 2007), although it was argued that gender differences in mathematical problem-solving resulted from stereotype threats depressing women's performance (Quinn & Spencer, 2001). In a science problem-solving study, male students were more successful because of their solution-seeking behaviour (Harskamp et al., 2008).

The present research

Our aim was to systematically investigate domain-specific rule knowledge and domain-general abilities in 6- to 8-year-old children's problem-solving. Therefore, we used three tablet-based *problem-solving tasks* in which participants had to solve construction problems with gears and building blocks according to well-defined goal requirements: (a) *carousel task*, (b) *propeller task* and (c) *stabilization task*. Additionally, three *rule knowledge tasks* assessed the domain-specific rule knowledge underlying the *problem-solving tasks*: (a) the *turning direction* of gears (rule: if adjacent connected gears turn, then their turning direction is opposite), (b) the *turning speed* of gears (rule: if a gear is larger than another gear, then it will turn comparably slower around its own axis) and (c) the stability of asymmetric building block constructions (rule: if the *centre-of-mass* of a building block construction is supported, it will remain stable). Furthermore, we measured two domain-general cognitive abilities: vocabulary as proxy for crystallized intelligence and reasoning as proxy for fluid intelligence. Our research questions were to test (1) whether problem-solving performance is more strongly related to (a) domain-specific rule knowledge or to (b) domain-general cognitive abilities (represented by crystallized and fluid intelligence and problem-solving in other domains; see Figure 2) and (2) whether age and gender have an impact on problem-solving performance.

METHODS

Participants

A power-analysis (parameters: $r = .25$, $\alpha = .05$, $1–\beta = .90$) resulted in a required sample size of $n = 164$. The sample consisted of 215 children (age: 6–8 years, M = 7.18, SD = 0.78; 89 female) from public schools. Data collection took place at a workshop during a school holiday at the university. They participated voluntarily and with written informed consent of their parents. The study was approved by the local ethics committee at the department of psychology (application #361). Ethnic background of participants was not recorded.

Design

We conducted a within-subject design, in which participants completed two sessions of approximately 30 min each with a 60-min break in between. Participants performed rule knowledge, vocabulary and reasoning tasks in the first session and problem-solving tasks (carousel, propeller and stabilization task) in the second session. In addition, children performed two tasks using analogue gear materials (cf. Schäfer et al., 2023a) in the second session, which are not subject of this study. Task order of the turning direction problem-solving task (carousel) and the turning speed problem-solving task (propeller) was counterbalanced, because they both belong to the gear domain. In relation to Bibi and Ahmad (2022), we presented non-routine problems since gears and building blocks are typically not used in goal-oriented problem-solving tasks, but as play materials. Trained experimenters explained and conducted the tasks in one-to-one settings with standardized procedures without providing feedback on response accuracy. All tasks were performed on a tablet application (cf. Schäfer et al., 2023a). Participants had sufficient time and instruction to get familiar with the task-specific tablet handling prior to each task. The procedure was video recorded.

Tests

The study comprised three rule knowledge tasks and three corresponding problem-solving tasks. All tasks started with one practice trial followed by the test trials (Figure 1).

Turning direction tasks

The *turning direction knowledge task* (Reuter & Leuchter, 2022b) contained six items with two to four connected gears of the same size. A driving gear was marked with a circular arrow indicating its turning direction. Participants were instructed to infer the turning direction(s) of the other connected gear(s) when the driving gear would turn in the direction specified by the arrow. Gears could neither be removed nor turned. Gears whose turning direction was estimated correctly were scored as one point each (0–13 points).

The *turning direction problem-solving task* (carousel task; Reuter & Leuchter, 2022b) contained a construction board with 20 × 14 plug-in options, and 14 gears of 4 different sizes arranged outside the board. The left lower corner of the board showed a fixed turnable driving gear and the upper right corner showed a fixed turnable target gear (i.e., a carousel). Participants were asked to build a connection between driving and target gear so that they turned in the same direction (goal state). Gears were moved via drag-and-drop and turned by circular swiping. Participants had up to 5 min to solve the task. Solution quality was ordinally scaled (0–4 points; see Table 1).

Turning speed tasks

The *turning speed knowledge task* (Reuter & Leuchter, 2022a) contained nine items with one to three connected gears of the same size in the upper half of the screen. In the lower half of the screen, three grey gears of different sizes were displayed. Participants had to pick the grey gear that would turn faster, slower or equally fast as the one(s) in the upper display. Gears could neither be removed nor turned. Gears whose turning speed was estimated correctly were scored as one point (0–9 points).

The *turning speed problem-solving task* (Propeller task; Reuter & Leuchter, 2019) contained a construction board with 20 × 14 plug-in options and 14 gears of 4 different sizes and 2 propellers

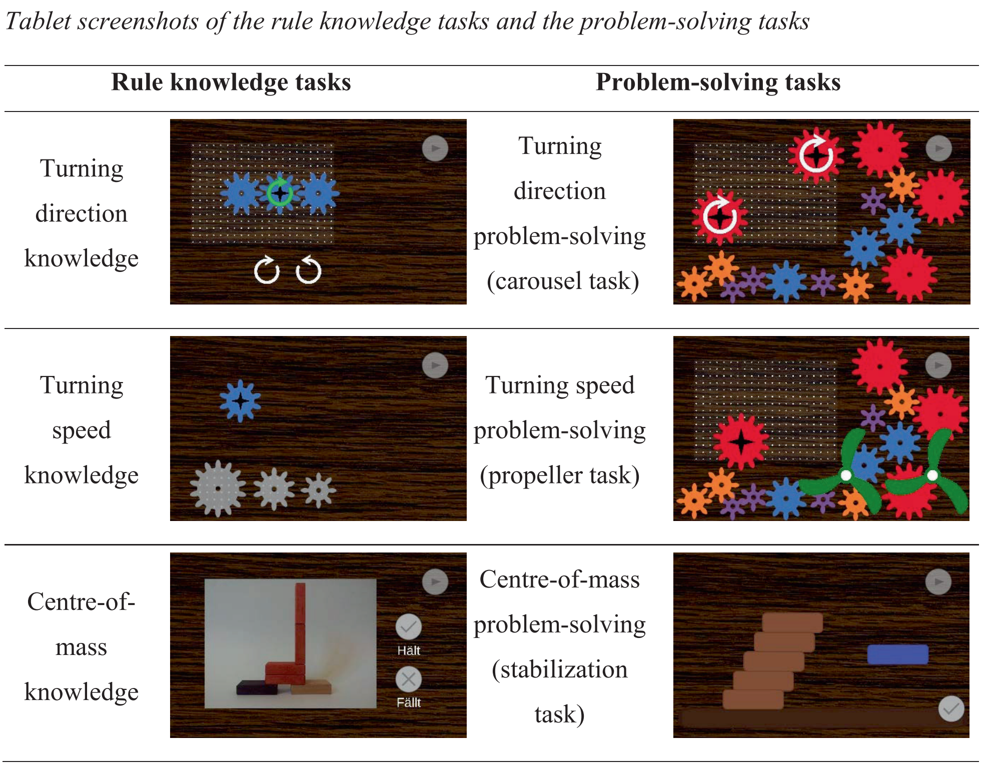

FIGURE 1 Tablet screenshots of the rule knowledge tasks and the problem-solving tasks.

TABLE 1 Grading system for the solution quality in the turning direction problem-solving task (carousel task).

Coded score	Final state on the gear board
0	Neither the driving gear nor the target gear had another gear attached to it
1	Either the driving gear or the target gear had at least one other gear attached to it
2	At least one gear was attached to both the driving and target gear
3	Driving gear and target gear were connected, but the target gear was either turning in the wrong direction or it was blocked
4	Driving gear and target gear were connected and the target gear turned in the specified direction when the driving gear turned in the specified direction

arranged outside the board. The lower part of the board showed a fixed turnable driving gear. The propellers could be attached to all gears except of the driving gear. Participants were asked to make one propeller turn as fast as possible and another propeller turn as slow as possible when turning the driving gear, without both propellers touching each other (goal state). Gears were moved via drag-and-drop and turned by circular swiping. Participants had up to 5 min to solve the task. Solution quality was ordinally scaled as two variables that were added to a sum score (turning speed 0–3 points plus contact 0–2 points; see Table 2).

Centre-of-mass tasks

The *centre-of-mass knowledge task* (Weber & Leuchter, 2020) contained 16 items of asymmetrical building block constructions. Participants rated whether a construction would hold or collapse

TABLE 2 Grading system for the solution quality in the turning speed problem-solving task (propeller task).

Coded score	Coding variable	Final state on the gear board
0	Turning speed	At least one propeller was not attached on any gear
1	Turning speed	Both propellers were attached on gears of the same size
2	Turning speed	Both propellers were attached on different-sized gears, but not largest and smallest
3	Turning speed	One propeller was attached on the smallest gear and the other propeller was attached on the largest gear
0	Contact	Not both propellers were driven by the driving gear or not both propellers were attached to the centre of a gear wheel
1	Contact	The two propellers touched each other when turning the driving gear
2	Contact	The two propellers did not touch each other when turning the driving gear

Note: The sum of the values in turning speed and in contact represent the total score.

after removing one specific colour-coded building block. Correctly solved items were scored as one point (0–16 points).

The *stabilization problem-solving task* (based on Weber et al., 2020) contained eight items of an instable asymmetrical building block construction, which could be stabilized (goal state) by placing an additional building block via drag-and-drop. Once the participants added the building block correctly or after three attempts, they were presented the next item. Each participant had a start budget of 32 points. For each item that was not successfully solved after three attempts and for each used attempt, one point was subtracted ($\text{Score}_{\text{stabilization}} = 32 - n_{\text{attempts}} - n_{\text{unresolvedItems}}$), resulting in 0–24 points.

Domain-general measures

We assessed vocabulary and reasoning as domain-general indicators for intelligence that are known to contribute to problem-solving performance.

Vocabulary: We used the passive vocabulary task (WPPSI-IV, Petermann & Daseking, 2018) as an indicator of crystallized intelligence since vocabulary enables the problem solver to verbally understand problem requirements and is closely related to acquired general mental abilities (Jensen, 2001). Participants saw four images on the display and were asked to point to the image named by the experimenter. The task consisted of up to 35 items. After three consecutive failures, the task was aborted. Correctly solved items were scored as one point (0–35 points).

Reasoning: We used the figural reasoning task (WPPSI-IV, Petermann & Daseking, 2018) as an indicator of fluid intelligence (Beauducel et al., 2001). It included 26 items with a 2×2 matrix containing three images and one question mark. Participants selected one of four to five figurative response options to replace the question mark in a logically consistent manner. After three consecutive failures, the task was aborted. Correctly solved items were scored as one point (0–26 points).

Statistical analyses

Descriptives and correlations (Spearman's Rho) were calculated for performance in all tasks. Correlations between corresponding rule knowledge and problem-solving tasks indicate domain-specific associations. Correlations between different problem-solving tasks indicate domain-general associations since they represent higher-level domain-general problem-solving abilities (see Figure 2). Correlations of problem-solving tasks with vocabulary and reasoning also indicate domain-general associations, *Model of research questions*

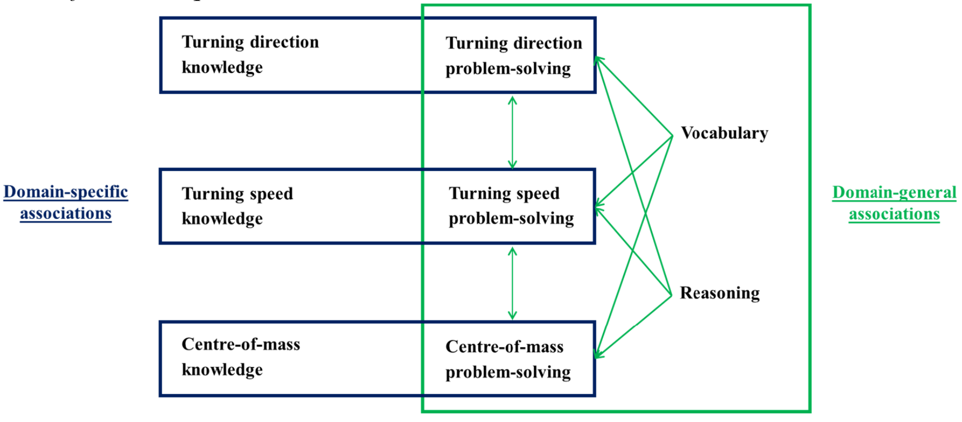

FIGURE 2 Model of research questions. Associations between tasks connected by a green arrow are domain-general. Associations of tasks in the same blue box are domain-specific.

because they may contribute to problem-solving across domains. Domain-specific and domain-general correlation strengths were compared according to Silver et al.'s (2004) method for correlations of independent groups.

The differential effects of domain-specific rule knowledge and domain-general abilities on problem-solving performance were analysed by a hierarchical ordinal logistic regression (proportional odds) for each problem-solving task as dependent variable (see Table 3). First, the domain-specific rule knowledge was the only included independent variable. Second, performances in other problem-solving tasks were added as independent variables. Third, vocabulary and reasoning were added as independent variables. Likelihood-ratio tasks were conducted to control whether the added variables increased the variance explained by the model. We excluded all participants with at least one missing task ($n = 12$). Additionally, we calculated confirmatory factor analyses based on structural equation models (see Appendix B (Figures A1–A3, Table A4)).

Age and gender effects on problem-solving performance were analysed through ordinal logistic regressions. To keep regression coefficients comparable, we z-standardized all independent factors, except of gender which was dichotomously dummy-coded. All analyses were performed using R statistics software (version 4.3.0; R Core Team, 2023).

RESULTS

Unifactorial ordinal regressions exhibited that task order neither had significant effects on turning direction problem-solving (carousel task: $\beta_{to} = -0.09$, $p = .47$) nor on turning speed problem-solving (propeller task: $\beta_{to} = 0.01$, $p = .93$).

Correlations and their comparisons

Correlations between both turning direction tasks and both turning speed tasks were significant and not significant between both centre-of-mass tasks. All correlations between different problem-solving tasks and all correlations of domain-general measures and problem-solving tasks were significant (Table 4).

The comparison of correlations showed that the association between corresponding domain-specific knowledge and problem-solving tasks (turning direction, turning speed, centre-of-mass) were

TABLE 3 Dependent and independent variables of regression models.

Model	Turning direction knowledge	Turning speed knowledge	Centre-of-mass knowledge	Reasoning	Vocabulary	Turning direction problem-solving	Turning speed problem-solving	Centre-of-mass problem-solving
Turning Direction	IV			IV	IV	DV	IV	IV
Turning Speed		IV		IV	IV	IV	DV	IV
Centre-of-mass			IV	IV	IV	IV	IV	DV

Abbreviations: DV, dependent variable; empty = not integrated in the model; IV, independent variable.

descriptively smaller than the correlations between the problem-solving tasks and other measures (i.e., other problem-solving tasks, vocabulary and reasoning). In 8 out of 12 possible combinations, this difference was significant (see Table 5).

Regression analyses

For all three problem-solving tasks (see Tables 6–8), second-step regression models accounted for additional variance in the respective problem-solving performance compared to first-step models. Regarding performance on the turning direction and turning speed problem-solving tasks, domain-specific rule knowledge was no longer a significant predictor once performance on the other problem-solving tasks was included as predictors. For centre-of-mass problem-solving, domain-specific rule knowledge was not a significant predictor. The third step models, adding vocabulary and reasoning, again accounted for additional variance. Reasoning was a significant predictor of the turning direction of problem-solving performance. Vocabulary was a significant predictor of the turning speed problem-solving performance and the centre-of-mass problem-solving performance (see also regression analyses with an intelligence composite score (Appendix A (Tables A1–A3)) and confirmatory factor analyses (Appendix B)).

Age and gender effects

Ordinal regressions showed that both age and gender were significant predictors of problem-solving performance across all tasks (see Table 9). The impact of age and gender was similarly strong (age: χ^2 (2) = 3.74–6.07; gender: χ^2 (2) = 2.50–4.05) and in favour of older and male participants.

DISCUSSION

The aim of this study was to test whether children's problem-solving performance was more strongly related to their domain-specific rule knowledge or to domain-general cognitive abilities. We also investigated the effects of age and gender on problem-solving performance.

Problem-solving performance was correlated with both domain-specific rule knowledge and domain-general cognitive abilities. Children's domain-specific rule knowledge about gear-turning mechanisms was correlated with performance in the corresponding problem-solving tasks, while rule knowledge about stability was not. These findings are partially in line with the expert-novice paradigm, that emphasizes the importance of content knowledge for problem-solving (e.g., Chi et al., 1981; Perkins & Salomon, 1989). Regarding domain-general cognitive abilities, vocabulary and reasoning as indicators of intelligence were correlated with all problem-solving tasks and explained variance in problem-solving performance beyond domain-specific knowledge. This is consistent with studies by Greiff et al. (2015) and Stephan et al. (2022) demonstrating substantial effects of intelligence on problem-solving. Correlational and regression analyses revealed comparably stronger associations

TABLE 4 Correlations (Spearman's Rho) of all tasks.

Variable	Range	M	SD	1	2	3	4	5	6	7
1. Turning direction knowledge	0–13	7.77	3.19							
2. Turning direction problem-solving (carousel task)	0–4	3.06	1.18	.20**						
3. Turning speed knowledge	0–9	5.29	2.92	.14*	.26**					
4. Turning speed problem-solving (propeller task)	0–5	2.43	1.66	.29**	.46**	.22**				
5. Centre-of-mass knowledge	0–16	5.85	3.10	.05	−.02	.02	−.07			
6. Centre-of-mass problem-solving (stabilization task)	0–24	12.22	3.20	.30**	.28**	.24**	.36**	.00		
7. Vocabulary	0–35	27.02	4.44	.33**	.34**	.22**	.38**	−.04	.37**	
8. Reasoning	0–26	16.20	4.47	.38**	.43**	.21**	.34**	.00	.36**	.41**

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

TABLE 5 Comparison of domain-specific and domain-general problem-solving correlation strengths.

Correlated variable	Turning direction problem-solving (carousel task) z (p)	Turning speed problem-solving (propeller task) z (p)	Centre-of-mass problem-solving (stabilization task) z (p)	Reasoning z (p)	Vocabulary z (p)	Reference
Turning direction problem-solving (carousel task)	–	3.47 (.00)	1.03 (.30)	3.24 (.00)	1.86 (.06)	Turning direction
Turning speed problem-solving (propeller task)	3.15 (.00)	–	1.76 (.08)	1.48 (.14)	2.00 (.04)	Turning speed
Centre-of-mass problem-solving (stabilization task)	2.92 (.00)	3.72 (.00)	–	3.85 (.00)	3.89 (.00)	Centre-of-mass

Note: Domain-specific correlations are compared with domain-general correlations. Positive z-values represent the domain-specific correlation to be weaker. *p*-Values indicate whether this difference is statistically significant.

TABLE 6 Ordinal logistic regression predicting performance on the turning direction problem-solving task (carousel task).

DV: Turning direction problem-solving performance (carousel task)			
	Step 1	Step 2	Step 3
Factor	Coef (SE)	Coef (SE)	Coef (SE)
Turning direction knowledge	.36** (.14)	.06 (.15)	−.11 (.16)
Turning speed problem-solving (propeller task)		.86** (.16)	.73** (.17)
Centre-of-mass problem-solving (stabilization task)		.21 (.15)	.09 (.16)
Vocabulary			.17 (.16)
Reasoning			.66** (.16)
Model parameters			
Log-likelihood	−247.54 (df = 5)	−227.25 (df = 7)	−217.19 (df = 9)
R^2	.04	.23	.31
Observations	203	203	203

$^*p < .05$. $^{**}p < .01$.

TABLE 7 Ordinal logistic regression predicting performance on the turning speed problem-solving task (propeller task).

DV: Turning speed problem-solving performance (propeller task)			
	Step 1	Step 2	Step 3
Factor	Coef (SE)	Coef (SE)	Coef (SE)
Turning speed knowledge	.36** (.13)	.15 (.13)	.12 (.13)
Turning direction problem-solving (carousel task)		.69** (.14)	.58** (.15)
Centre-of-mass problem-solving (stabilization task)		.48** (.13)	.36** (.14)
Vocabulary			.38** (.14)
Reasoning			.19 (.15)
Model parameters			
Log-Likelihood	−348.64 (df = 6)	−325.51 (df = 8)	−319.91 (df = 10)
R^2	.04	.24	.29
Observations	203	203	203

$^*p < .05$. $^{**}p < .01$.

TABLE 8 Ordinal logistic regression predicting performance on the centre-of-mass problem-solving task (stabilization task).

DV: Centre-of-mass problem-solving (stabilization task)			
	Step 1	Step 2	Step 3
Factor	Coef (SE)	Coef (SE)	Coef (SE)
Centre-of-mass knowledge	.03 (.12)	.09 (.12)	.10 (.13)
Turning direction problem-solving (carousel task)		.31* (.14)	.19 (.14)
Turning speed problem-solving (propeller task)		.52** (.14)	.37** (.14)
Vocabulary			.36* (.14)
Reasoning			.28 (.14)
Model parameters			
Log-likelihood	−510.11 (df = 16)	−494.48 (df = 18)	−487.63 (df = 20)
R^2	.00	.14	.20
Observations	203	203	203

$^*p < .05$. $^{**}p < .01$.

TABLE 9 Ordinal logistic regression predicting performance on problem-solving tasks by age and gender.

	Turning direction problem-solving (carousel task)	Turning speed problem-solving (propeller task)	Centre-of-mass problem-solving (stabilization task)
Factor	Coef (SE)	Coef (SE)	Coef (SE)
Age	.53** (.14)	.81** (.14)	.84** (.14)
Gender	.69* (.28)	1.08** (.27)	.78** (.25)
Model parameters			
Log-likelihood	−239.96 (df = 6)	−324.69 (df = 7)	−484.95 (df = 17)
R^2	.11	.25	.22
Observations	203	203	203

$^*p < .05$. $^{**}p < .01$.

between performance in different problem-solving tasks than between domain-specific knowledge and corresponding problem-solving performance. Moreover, comparisons of correlation strengths showed that performance in one specific problem-solving task was more strongly related to the performance in the other problem-solving tasks and to domain-general cognitive abilities than to the corresponding domain-specific rule knowledge. Additionally, for all problem-solving tasks (carousel, propeller, stabilization), effects of rule knowledge on performance became non-significant when including the other two problem-solving measures as independent variables in the regression analyses.

Taken together, our findings indicate that domain-specific rule knowledge played a minor role for 6- to 8-year-olds' science problem-solving performance, at least compared to domain-general abilities, such as intelligence and general problem-solving abilities. This finding is supported by the results of Geary et al. (2017), who found that domain-general cognitive abilities were more important than domain-specific knowledge for mathematics problem-solving in children. In contrast to our findings, English (1992) and Gilmore et al. (2018) found that content knowledge played a major role in combinatorial and arithmetic problem-solving. However, the relevance of domain-specific knowledge might increase as a function of age (cf. Geary et al., 2017).

We also found that age was a significant predictor of performance on all problem-solving tasks. Older children outperformed younger children, which is in line with previous research (Keen, 2011; Reuter & Leuchter, 2021; Thornton, 2009). Moreover, male children outperformed female ones on the problem-solving tasks. This finding adds to previous research that pointed to gender differences in science problem-solving performance (e.g., Harskamp et al., 2008) and may be of importance for the development of STEM curricula that should provide equal learning opportunities for all children (Shechter et al., 2021). Further studies will be necessary to clarify the role of context factors, such as play materials available for girls and boys, and of task modality, such as spatial compared to verbal or figural tasks.

Taken together, the findings suggest that domain-specific rule knowledge plays a minor role for problem-solving performance in 6- to 8-year-old children, compared to domain-general abilities, such as vocabulary, reasoning and general problem-solving abilities. Two assumptions might explain these results.

First, children might have solved the problems through trial-and-error irrespective of whether they had rule knowledge or not, so that rule knowledge would have been irrelevant (Funke et al., 2018; Tönnsen, 2021). Trial-and-error might be helpful for children without domain-specific rule knowledge and for children who are incapable to retrieve it (Sabella & Redish, 2007; Tuminaro & Redish, 2007). Furthermore, the environment of our problem-solving tasks was playful and might thus have led children to try different solution pathways by trial-and-error, because they assumed that their performance on the tasks was not relevant. However, participants who did not have solid domain-specific rule knowledge prior to problem-solving might have acquired this knowledge during the problem-solving process (Tönnsen, 2021). In particular, trial-and-error might integrate implicit rule knowledge which is, however, hardly assessable (Funke et al., 2018). To gain more insights into whether rule knowledge is acquired during trial-and-error, microgenetic studies (cf. Siegler, 2007) are a promising approach. Moreover, the relevance of domain-specific knowledge may depend on individual problem-solving strategies. Thus, applied strategies should be analysed in future studies, for example based on number and timing of gear and block manipulations.

Second, children who have domain-specific rule knowledge might fail to apply it. This failure might depend on more general cognitive skills, for example executive functioning (Dixon & Johnson, 2012; Schäfer et al., 2023b), which continuously develop till late adolescence (Wiebe & Karbach, 2017). Children's limited working memory capacity might affect their problem-solving performance, because the problems required the maintenance of the objects' continuously changing constellations, such as the interrelations of the gears or of the building blocks. Deficits in executive functioning might also affect the capability to remember previous errors systematically, leading children to commit the same errors repeatedly, thus hindering learning from errors. This explanation is in line with the finding that working memory capacity mediates implicit learning in problem-solving (Reber & Kotovsky, 1997).

Limitations and future research

Although we gained new insights into children's problem-solving, the rule knowledge tasks applied may not fully capture all necessary aspects of the respective domain-specific knowledge. We tested domain-specific rule knowledge and corresponding problem-solving in only two domains (gears and building blocks), revealing a minor role of domain-specific rule knowledge. However, the prominent role of domain-specific knowledge in problem-solving, observed in expert-novice studies (Perkins & Salomon, 1989), might be latently inherent at age 6–8 years.

To study the development of relations between domain-specific rule knowledge, domain-general cognitive abilities and problem-solving, a longitudinal design is promising. Generalizing our findings requires similar research in other domains. Assessing rule knowledge before and after problem-solving could illuminate the acquisition of rule knowledge during problem-solving.

Our results highlight vocabulary and reasoning as central domain-general aspects of problem-solving (cf. Greiff et al., 2015). However, given that Bühner et al. (2008) found effects of working memory on problem-solving abilities, the assessment of a broader set of cognitive abilities would allow more specific conclusions. Verbal presentation of our problem-solving tasks, requiring children to maintain instructions for several minutes, may be affected by intelligence and working memory.

Additionally, given the spatial representation and goal-oriented nature of our tasks, spatial abilities and inhibition are worth further investigation (Dixon & Johnson, 2012; Oostermeijer et al., 2014).

CONCLUSION

Our study indicates a major role of domain-general cognitive abilities in 6- to 8-year-old children's science problem-solving. We found that domain-general cognitive abilities had a stronger association with problem-solving abilities in three different problems than domain-specific rule knowledge. This suggests the importance of studying problem-solving even in participants lacking substantial domain-specific rule knowledge, thus extending beyond the expert-novice paradigm. Regarding educational settings, assessing problem-solving in the classroom can be valuable, especially when problem-solving is part of the curriculum.

AUTHOR CONTRIBUTIONS

Jonas Schäfer: Conceptualization; data curation; formal analysis; investigation; methodology; visualization; writing – original draft; writing – review and editing; software; validation. **Timo Reuter:** Conceptualization; investigation; writing – review and editing; methodology. **Julia Karbach:** Conceptualization; resources; writing – review and editing. **Miriam Leuchter:** Conceptualization; funding acquisition; project administration; resources; supervision; writing – review and editing.

ACKNOWLEDGEMENTS

The authors would like to thank Steffen Pohl for his excellent support in programming of experimental software. We further want to thank the student research assistants and all the participating children for realizing the data collection. Open Access funding enabled and organized by Projekt DEAL.

FUNDING INFORMATION

This research was supported by the Ministry of Science, Further Education, Research and Culture, Rhineland-Palatinate, Germany.

CONFLICT OF INTEREST STATEMENT

The authors do not report any conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Jonas Schäfer https://orcid.org/0000-0002-3019-8236
Timo Reuter https://orcid.org/0000-0002-2098-206X
Julia Karbach https://orcid.org/0000-0003-2763-6351
Miriam Leuchter https://orcid.org/0000-0002-7962-6561

REFERENCES

Adams, R. S., & Atman, C. J. (1999). *Cognitive processes in iterative design behavior.* FIE'99 Frontiers in Education. 29th Annual Frontiers in Education Conference. Designing the Future of Science and Engineering Education. Conference Proceedings (*IEEE Cat. No.99CH37011*, San Juan, PR, USA, 1999, vol. 1, pp. 11A6/13-11A6/18). https://doi.org/10.1109/FIE.1999.839114

Alexander, P. A., & Judy, J. E. (1988). The interaction of domain-specific and strategic knowledge in academic performance. *Review of Educational Research*, *58*(4), 375–404. https://doi.org/10.3102/00346543058004375

Arló-Costa, H., & Levi, I. (1996). Two notions of epistemic validity. *Synthese*, *109*, 217–262. https://doi.org/10.1007/BF00413768

Beauducel, A., Brocke, B., & Liepmann, D. (2001). Perspectives on fluid and crystallized intelligence: Facets for verbal, numerical, and figural intelligence. *Personality and Individual Differences*, *30*(6), 977–994. https://doi.org/10.1016/S0191-8869(00)00087-8

Beauducel, A., & Wittmann, W. W. (2005). Simulation study on fit indices in confirmatory factor analysis based on data with slightly distorted simple structure. *Structural Equation Modeling*, *12*, 41–75. https://doi.org/10.1207/s15328007sem1201_3

Bibi, A., & Ahmad, M. (2022). Goal orientations, self-regulated learning strategies and problem-solving: A mediational analysis. *Pertanika Journal of Social Sciences and Humanities*, *30*(3), 949–976. https://doi.org/10.47836/pjssh.30.3.02

Bibi, A., Syed-Zamri, S. N. A. B., & Abedalaziz, N. A. M. (2018). Role of students' context familiarity in differential equations problem solving at pre-university level. *Malaysian Online Journal of Educational Sciences*, *6*(4), 47–57. https://doi.org/10.22342/jme.v13i2.pp323-336

Bühner, M., Kröner, S., & Ziegler, M. (2008). Working memory, visual-spatial intelligence and their relationship to problem solving. *Intelligence*, *36*(6), 672–680. https://doi.org/10.1016/j.intell.2008.03.008

Charlesworth, R., & Leali, S. A. (2012). Using problem solving to assess young children's mathematics knowledge. *Early Childhood Education Journal*, *39*, 373–382. https://doi.org/10.1007/s10643-011-0480-y

Chen, X., Hertzog, C., & Park, D. C. (2017). Cognitive predictors of everyday problem solving across the lifespan. *Gerontology*, *63*(4), 372–384. https://doi.org/10.1159/000459622

Chi, M. T. (2006). Laboratory methods for assessing experts' and novices' knowledge. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hofmann (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (pp. 167–184). Cambridge University Press. https://doi.org/10.1017/CBO9780511816796.010

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*, 121–152. https://doi.org/10.1207/s15516709cog0502_2

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. *1*). Erlbaum.

Chiappe, D., & MacDonald, K. (2005). The evolution of domain-general mechanisms in intelligence and learning. *The Journal of General Psychology*, *132*(1), 5–40. https://doi.org/10.3200/GENP.132.1.5-40

Diamond, L. L. (2018). Problem solving in the early years. *Intervention in School and Clinic*, *53*(4), 220–223. https://doi.org/10.1177/1053451217712957

diSessa, A. A. (1988). Knowledge in pieces. In G. Forman & P. Pufall (Eds.), *Constructivism in the computer age*. Lawrence Erlbaum Publishers.

Dixon, A. R., & Johnson, S. D. (2012). The use of executive control processes in engineering design by engineering students and professional engineers. *Journal of Technology Education*, *24*(1), 73–89. https://doi.org/10.21061/jte.v24i1.a.5

English, L. (1992). Children's use of domain-specific knowledge and domain-general strategies in novel problem solving. *British Journal of Educational Psychology*, *62*(2), 203–216. https://doi.org/10.1111/j.2044-8279.1992.tb01014.x

Fischer, A., Greiff, S., & Funke, J. (2011). The process of solving complex problems. *The Journal of Problem Solving*, *4*(1), 19–42.

Fitzgerald, A., & Smith, K. (2016). Science that matters: Exploring science learning and teaching in primary schools. *Australian Journal of Teacher Education*, *41*(4), 64–78. https://doi.org/10.14221/ajte.2016v41n4.4

Funke, J. (2014). Analysis of minimal complex systems and complex problem solving require different forms of causal cognition. *Frontiers in Psychology*, *5*, 739. https://doi.org/10.3389/fpsyg.2014.00739

Funke, J., Fischer, A., & Holt, D. V. (2018). Competencies for complexity: Problem solving in the twenty-first century. In E. Care, P. Griffin, & M. Wilson (Eds.), *Educational assessment in an information age. Assessment and teaching of 21st century skills* (pp. 41–53). Springer International Publishing. https://doi.org/10.1007/978-3-319-65368-6_3

Geary, D. C., Nicholas, A., Li, Y., & Sun, J. (2017). Developmental change in the influence of domain-general abilities and domain-specific knowledge on mathematics achievement: An eight-year longitudinal study. *Journal of Educational Psychology*, *109*(5), 680–693. https://doi.org/10.1037/edu0000159

Gilmore, C., Clayton, S., Cragg, L., McKeaveney, C., Simms, V., & Johnson, S. (2018). Understanding arithmetic concepts: The role of domain-specific and domain-general skills. *PLoS One*, *13*(9), e0201724. https://doi.org/10.1371/journal.pone.0201724

Gold, Z. S., & Elicker, J. (2020). Engineering peer play: A new perspective on science, technology, engineering, and mathematics (STEM) early childhood education. In A. Ridgway, G. Quiñones, & L. Li (Eds.), *Peer play and relationships in early childhood. International perspectives on early childhood education and development* (p. 30). Springer. https://doi.org/10.1007/978-3-030-42331-5_5

Greene, J. A., Cartiff, B. M., & Duke, R. F. (2018). A meta-analytic review of the relationship between epistemic cognition and academic achievement. *Journal of Educational Psychology*, *110*(8), 1084–1111. https://doi.org/10.1037/edu0000263

Greiff, S., Krkovic, K., & Hautamäki, J. (2015). The prediction of problem solving assessed via microworlds. *European Journal of Psychological Assessment*, *32*(4), 298–306. https://doi.org/10.1027/1015-5759/a000263

Greiff, S., Wüstenberg, S., Csapó, B., Demetriou, A., Hautamäki, J., Graesser, A. C., & Martin, R. (2014). Domain-general problem solving skills and education in the 21st century. *Educational Research Review*, *13*, 74–83. https://doi.org/10.1016/j.edurev.2014.10.002

Gunzenhauser, C., Karbach, J., & Saalbach, H. (2019). Function of verbal strategies in monolingual vs. bilingual students' planning performance: An experimental approach. *Cognitive Development*, *50*, 1–12. https://doi.org/10.1016/j.cogdev.2019.01.003

Hamilton, E., Lesh, R., Lester, F., & Yoon, C. (2007). The use of reflection tools to build personal models of problem-solving. In A. R. Lesh, E. Hamilton, & J. J. Kaput (Eds.), *Foundations for the future in mathematics education*. Routledge. https://doi.org/10.4324/9781003064527-21

Harskamp, E., Ding, N., & Suhre, C. (2008). Group composition and its effect on female and male problem-solving in science education. *Educational Research*, *50*(4), 307–318. https://doi.org/10.1080/00131880802499688

Hayduk, A. L., & Littvay, L. (2012). Should researchers use single indicators, best indicators, or multiple indicators in structural equation models? *BMC Medical Research Methodology*, *12*(1), 159–175. https://doi.org/10.1186/1471-2288-12-159

Injoque-Ricle, I., Barreyro, J. P., Calero, A., & Burin, D. I. (2014). Tower of London: Planning development in children from 6 to 13 years of age. *The Spanish Journal of Psychology*, *17*, E77. https://doi.org/10.1017/sjp.2014.83

Jensen, A. R. (2001). Vocabulary and general intelligence. *The Behavioral and Brain Sciences*, *24*(6), 1109–1110. https://doi.org/10.1017/S0140525X01280133

Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development*, *48*, 63–85. https://doi.org/10.1007/BF02300500

Keen, R. (2011). The development of problem solving in young children: A critical cognitive skill. *Annual Review of Psychology*, *62*, 1–21. https://doi.org/10.1146/annurev.psych.031809.130730

Krist, H. (2010). Development of intuitions about support beyond infancy. *Developmental Psychology*, *46*(1), 266–278. https://doi.org/10.1037/a0018040

Krist, H., Horz, H., & Schönfeld, T. (2005). Children's block balancing revisited: No evidence for representational redescription. *Swiss Journal of Psychology*, *64*(3), 183–193. https://doi.org/10.1024/1421-0185.64.3.183

Lee, N. Y. L., & Johnson-Laird, P. N. (2013). Strategic changes in problem solving. *Journal of Cognitive Psychology*, *25*(2), 165–173. https://doi.org/10.1080/20445911.2012.719021

Lehrer, R., & Schauble, L. (1998). Reasoning about structure and function: Children's conceptions of gears. *Journal of Research in Science Teaching*, *35*(1), 3–25. https://doi.org/10.1002/(sici)1098-2736(199801)35:1<3::aid-tea2>3.0.co;2-x

Lucas, B., & Hanson, J. (2016). Thinking like an engineer: Using engineering habits of mind and signature pedagogies to redesign engineering education. *International Journal of Engineering Pedagogy (IJEP)*, *6*(2), 4–13. https://doi.org/10.3991/ijep.v6i2.5366

Luo, J., & Niki, K. (2003). Function of hippocampus in "insight" of problem solving. *Hippocampus*, *13*(3), 316–323. https://doi.org/10.1002/hipo.10069

Mayer, R. E. (1997). Chapter 25 - from novice to expert. In M. G. Helander, T. K. Landauer, & P. V. Prabhu (Eds.), *Handbook of human-computer interaction* (Second ed., pp. 781–795). North-Holland. https://doi.org/10.1016/B978-0-444-70536-5.50030-0

Mehadi Rahman, M. (2019). 21st century skill "problem solving": Defining the concept. *Asian Journal of Interdisciplinary Research*, *2*(1), 64–74. https://doi.org/10.34256/ajir1917

Milbourne, J., & Wiebe, E. (2018). The role of content knowledge in ill-structured problem solving for high school physics students. *Research in Science Education*, *48*, 165–179. https://doi.org/10.1007/s11165-016-9564-4

Mustakim, M., Mansyur, J., Hatibe, A., Rizal, M., & Kaharu, S. N. (2020). Analysis of Students' causal reasoning in physics problem solving. *Journal of Physics: Conference Series*, *1521*, 22058. https://doi.org/10.1088/1742-6596/1521/2/022058

O'Brien, S., Mitchell, D. J., Duncan, J., & Holmes, J. (2023). Cognitive segmentation and fluid reasoning in childhood. *The Quarterly Journal of Experimental Psychology*, *76*(6), 1431–1444. https://doi.org/10.1177/17470218221116054

Oostermeijer, M., Boonen, A. J. H., & Jolles, J. (2014). The relation between children's constructive play activities, spatial ability, and mathematical word problem solving performance: A mediation analysis in sixth-grade students. *Frontiers in Psychology*, *5*, 782. https://doi.org/10.3389/fpsyg.2014.00782

Perkins, D. N., & Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher*, *18*(1), 16–25. https://doi.org/10.3102/0013189X018001016

Petermann, F., & Daseking, M. (2018). *Wechsler preschool and primary scale of intelligence* (Fourth ed.). Hogrefe.

Quinn, D. M., & Spencer, S. J. (2001). The interference of stereotype threat with women's generation of mathematical problem solving strategies. *Journal of Social Issues*, *57*(1), 55–71. https://doi.org/10.1111/0022-4537.00201

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Austria. https://www.R-project.org/

Reber, P. J., & Kotovsky, K. (1997). Implicit learning in problem solving: The role of working memory capacity. *Journal of Experimental Psychology: General*, *126*(2), 178–203. https://doi.org/10.1037/0096-3445.126.2.178

Reif, F., & Heller, J. I. (1982). Knowledge structure and problem solving in physics. *Educational Psychologist*, *17*(2), 102–127. https://doi.org/10.1080/00461528209529248

Reuter, T., & Leuchter, M. (2019). *Preschooler's concepts and problem solving in the domain of gears in the context of guided play*. Presentation at the 18th biennial EARLI Conference. Aachen, August 2019.

Reuter, T., & Leuchter, M. (2021). Children's concepts of gears and their promotion through play. *Journal of Research in Science Teaching*, *58*(1), 69–94. https://doi.org/10.1002/tea.21647

Reuter, T., & Leuchter, M. (2022a). Examining kindergarten children's testing and optimising in the context of a gear engineering task. *European Journal of STEM Education*, *7*(1), 04. https://doi.org/10.20897/ejsteme/11827

Reuter, T., & Leuchter, M. (2022b). Fostering 5- to 6-year-old children's conceptual knowledge of gear functioning through guided play in German kindergartens. In S. D. Tunnicliffe & T. J. Kennedy (Eds.), *Play and STEM education in the early years: International policies and practices* (pp. 331–343). Cham.

Sabella, M. S., & Redish, E. F. (2007). Knowledge organization and activation in physics problem solving. *American Journal of Physics*, *75*, 1017–1029. https://doi.org/10.1119/1.2746359

Schäfer, J., Reuter, T., Leuchter, M., & Karbach, J. (2023a). Validation of new tablet-based problem-solving tasks in primary school students. In press.

Schäfer, J., Reuter, T., Leuchter, M., & Karbach, J. (2023b). Executive functions and problem-solving – The contribution of inhibition, working memory, and cognitive flexibility to science problem-solving performance in elementary school students. In press.

Schoenfeld, A. H. (1987). What's all the fuss about metacognition? In A. H. Schoenfeld (Ed.), *Cognitive science and mathematics education* (pp. 189–215). Erlbaum. https://doi.org/10.4324/9780203062685

Shechter, T., Eden, S., & Spektor-Levy, O. (2021). Preschoolers' nascent engineering thinking during a construction task. *Journal of Cognitive Education and Psychology*, *20*(2), 83–111. https://doi.org/10.1891/JCEP-D-20-00010

Siegler, R. S. (2007). Microgenetic analyses of learning. In D. Kuhn, R. S. Siegler, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Vol 2, cognition, perception, and language* (6th ed., pp. 464–510). John Wiley & Sons Inc.. https://doi.org/10.1002/9780470147658.chpsy0211

Silver, N. C., Hittner, J. B., & May, K. (2004). Testing dependent correlations with nonoverlapping variables: A Monte Carlo simulation. *The Journal of Experimental Education*, *73*(1), 53–69. https://doi.org/10.3200/JEXE.71.1.53-70

Simmons, R. G. (1992). The roles of associational and causal reasoning in problem solving. *Artificial Intelligence*, *53*(2), 159–207. https://doi.org/10.1016/0004-3702(92)90070-E

Spektor-Levy, O., & Shechter, T. (2022). Learning environments that improve STEM capabilities in Israel: Constructional play and Preschoolers' engineering habits of mind. In S. D. Tunnicliffe & T. J. Kennedy (Eds.), *Play and STEM education in the early years* (pp. 311–329). Springer Cham. https://doi.org/10.1007/978-3-030-99830-1_15

Stephan, F., Gunzenhauser, C., & Saalbach, H. (2022). Function of language skills in Preschooler's problem solving performance: The role of self-directed speech. *Journal of Applied Developmental Psychology*, *81*, 101431. https://doi.org/10.1016/j.appdev.2022.101431

Stern, M., & Hertel, S. (2022). Relationship between maternal scaffolding and Preschooler's metacognitive strategies in a problem-solving situation. *Learning and Instruction*, *80*, 101631. https://doi.org/10.1016/j.learninstruc.2022.101631

Strimel, G. J., Bartholomew, S. R., Kim, E., & Zhang, L. (2018). An investigation of engineering design cognition and achievement in primary school. *Journal for STEM Education Research*, *1*, 173–201. https://doi.org/10.1007/s41979-018-0008-0

Thornton, S. (2009). *Children solving problems*. Harvard University Press. https://doi.org/10.4159/9780674044340

Tönnsen, K.-C. (2021). The relevance of trial-and-error: Can trial-and-error be a sufficient learning method in technical problem-solving-contexts? *Techne Serien – Forskning i slöjdpedagogik Och slöjdvetenskap*, *28*(2), 303–312.

Tuminaro, J., & Redish, E. F. (2007). Elements of a cognitive model of physics problem solving: Epistemic games. *Physical Review Physics Education Research*, *3*(2), 020201-1–020201-22. https://doi.org/10.1103/PhysRevSTPER.3.020101

Walker, S., Irving, K., & Berthelsen, D. (2002). Gender influences on preschool children's social problem solving strategies. *The Journal of Genetic Psychology*, *163*(2), 197–209. https://doi.org/10.1080/00221320209598677

Weber, A. M., & Leuchter, M. (2020). Measuring preschool children's knowledge of the principle of static equilibrium in the context of building blocks: Validation of a test instrument. *The British Journal of Educational Psychology*, *90*, 50–74. https://doi.org/10.1111/bjep.12304

Weber, A. M., Reuter, T., & Leuchter, M. (2020). The impact of a construction play on 5- to 6-year-old children's reasoning about stability. *Frontiers in Psychology*, *11*, 1737. https://doi.org/10.3389/fpsyg.2020.01737

Weisberg, D. S., Hirsh-Pasek, K., & Michnick, G. R. (2013). Guided play. Where curricular goals meet a playful pedagogy. *Mind, Brain, and Education*, *7*(2), 104–112. https://doi.org/10.1111/mbe.12015

Wiebe, S. A., & Karbach, J. (Eds.). (2017). *Executive function: Development across the life span* (1st ed.). Routledge. https://doi.org/10.4324/9781315160719

Zhu, Z. (2007). Gender differences in mathematical problem solving patterns: A review of literature. *International Education Journal*, *8*(2), 187–203.

Zook, N. A., Davalos, D. B., Delosh, E. L., & Davis, H. P. (2004). Working memory, inhibition, and fluid intelligence as predictors of performance on tower of Hanoi and London tasks. *Brain and Cognition*, *56*(3), 286–292. https://doi.org/10.1016/j.bandc.2004.07.003

How to cite this article: Schäfer, J., Reuter, T., Karbach, J., & Leuchter, M. (2024). Domain-specific knowledge and domain-general abilities in children's science problem-solving. *British Journal of Educational Psychology*, *94*, 346–366. https://doi.org/10.1111/bjep.12649

APPENDIX A

REGRESSION ANALYSES WITH A COMPOSITE FACTOR OF INTELLIGENCE

TABLE A1 Ordinal logistic regression predicting performance on the turning direction problem-solving task (carousel task).

DV: Turning direction problem-solving performance (carousel task)			
	Step 1	Step 2	Step 3
Factor	Coef (SE)	Coef (SE)	Coef (SE)
Turning direction knowledge	.36** (.14)	.06 (.15)	−.08 (.16)
Turning speed problem-solving (propeller task)		.86** (.16)	.71** (.16)
Centre-of-mass problem-solving (stabilization task)		.21 (.15)	.08 (.16)
Intelligence			.40** (.10)
Model parameters			
Log-likelihood	−247.54 (df = 5)	−227.25 (df = 7)	−219.28 (df = 8)
R^2	.04	.23	.29
Observations	203	203	203

Note: The factor Intelligence represents the sum of the standardized scores of vocabulary and reasoning.
*p < .05. **p < .01.

TABLE A2 Ordinal logistic regression predicting performance on the turning speed problem-solving task (propeller task).

DV: Turning speed problem-solving performance (propeller task)			
	Step 1	Step 2	Step 3
Factor	Coef (SE)	Coef (SE)	Coef (SE)
Turning speed knowledge	.36** (.13)	.15 (.13)	.12 (.13)
Turning direction problem-solving (carousel task)		.69** (.14)	.56** (.14)
Centre-of-mass problem-solving (stabilization task)		.48** (.13)	.36** (.14)
Intelligence			.29** (.09)
Model parameters			
Log-likelihood	−348.64 (df = 6)	−325.51 (df = 8)	−320.25 (df = 9)
R^2	.04	.24	.29
Observations	203	203	203

Note: The factor Intelligence represents the sum of the standardized scores of vocabulary and reasoning.
*p < .05. **p < .01.

TABLE A3 Ordinal logistic regression predicting performance on the centre-of-mass problem-solving task (stabilization task).

DV: Centre-of-mass problem-solving (stabilization task)			
	Step 1	Step 2	Step 3
Factor	Coef (SE)	Coef (SE)	Coef (SE)
Centre-of-mass knowledge	.03 (.12)	.09 (.12)	.09 (.13)
Turning direction problem-solving (carousel task)		.31* (.14)	.19 (.14)
Turning speed problem-solving (propeller task)		.52** (.14)	.38** (.14)
Intelligence			.32** (.09)
Model parameters			
Log-likelihood	−510.11 (df = 16)	−494.48 (df = 18)	−487.69 (df = 19)
R^2	.00	.14	.20
Observations	203	203	203

Note: The factor Intelligence represents the sum of the standardized scores of vocabulary and reasoning.
*p < .05. **p < .01.

APPENDIX B

CONFIRMATORY FACTOR ANALYSES

Confirmatory factor analysis for the performance in turning direction problem-solving (carousel task)

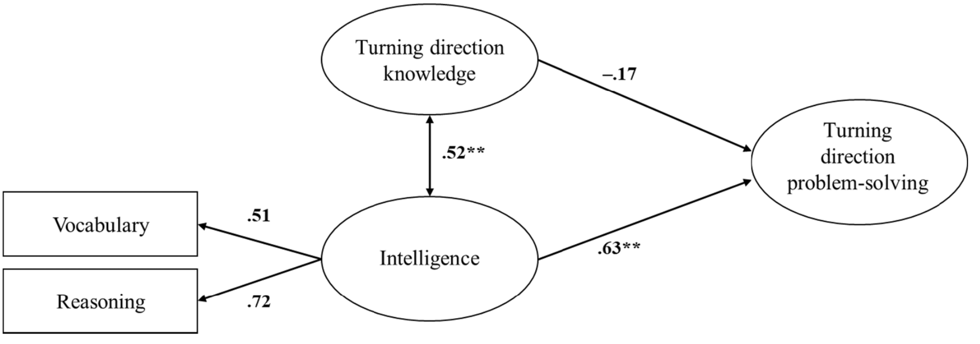

FIGURE A1 Confirmatory factor analysis for the performance in turning direction problem-solving (carousel task). *p < .05; **p < .01; Unidirectional arrows between latent factors represent regressions, bidirectional arrows represent correlations. Path values represent the standardized estimates. Domain-specific knowledge and problem-solving are defined as single-indicator latent factors (Hayduk & Littvay, 2012).

Confirmatory factor analysis for the performance in turning speed problem-solving

(propeller task)

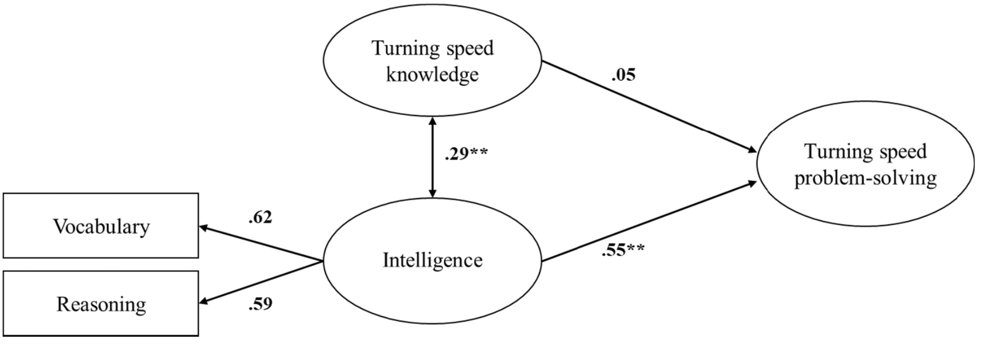

FIGURE A2 Confirmatory factor analysis for the performance in turning speed problem-solving (*propeller task*). *$p < .05$; **$p < .01$; Unidirectional arrows between latent factors represent regressions, bidirectional arrows represent correlations. Path values represent the standardized estimates. Domain-specific knowledge and problem-solving are defined as single-indicator latent factors (Hayduk & Littvay, 2012).

Confirmatory factor analysis for the performance in turning direction problem-solving

(stabilization task)

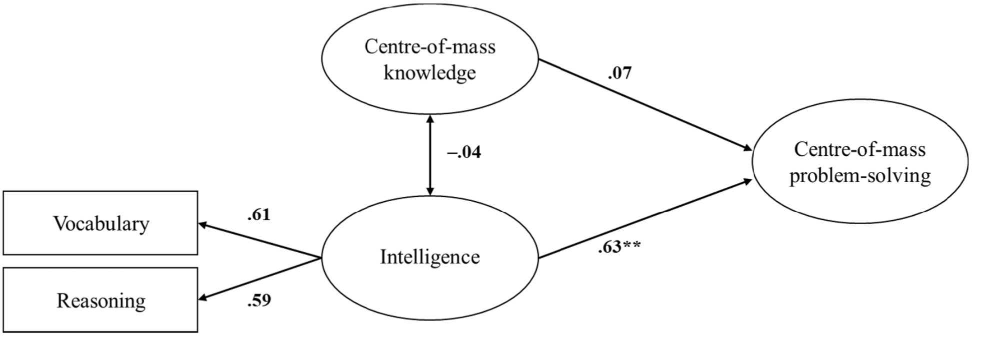

FIGURE A3 Confirmatory factor analysis for the performance in turning direction problem-solving (stabilization task). *$p < .05$; **$p < .01$; Unidirectional arrows between latent factors represent regressions, bidirectional arrows represent correlations. Path values represent the standardized estimates. Domain-specific knowledge and problem-solving are defined as single-indicator latent factors (Hayduk & Littvay, 2012).

TABLE A4 Model fit indices of the confirmatory factor analyses.

Model	χ^2 (df)	p	CFI	TLI	RMSEA (90% CI)	SRMR
Turning direction problem-solving (carousel task)	.26 (1)	.61	1.00	1.04	.00 (.00–.15)	.01
Turning speed problem-solving (propeller task)	.01 (1)	.94	1.00	1.08	.00 (.00–.05)	.00
Centre-of-mass problem-solving (stabilization task)	.18 (1)	.18	.99	.92	.06 (.00–.21)	.02

Note: Model fit indices are evaluated according to Beauducel and Wittmann (2005).

7.3 Study 3: Executive functions and problem-solving – the contribution of inhibition, working memory, and cognitive flexibility to science problem-solving performance in elementary school students

Schäfer, J., Reuter, T., Leuchter, M., & Karbach, J. (2024). Executive functions and problem-solving – the contribution of inhibition, working memory, and cognitive flexibility to science problem-solving performance in elementary school students. *Journal of Experimental Child Psychology, 244*, 105962. https://doi.org/10.1016/j.jecp.2024.105962

Journal of Experimental Child Psychology 244 (2024) 105962

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Executive functions and problem-solving—The contribution of inhibition, working memory, and cognitive flexibility to science problem-solving performance in elementary school students

Jonas Schäfer [a,b,*], Timo Reuter [b], Miriam Leuchter [b], Julia Karbach [a,c]

[a] Department of Psychology, University of Kaiserslautern–Landau, 76829 Landau, Germany
[b] Institute for Child and Youth Education, University of Kaiserslautern–Landau, 76829 Landau, Germany
[c] Center for Research on Individual Development and Adaptive Education of Children at Risk (IDeA), 60323 Frankfurt am Main, Germany

ARTICLE INFO

Article history:
Received 7 October 2023
Revised 16 April 2024
Accepted 17 April 2024

Keywords:
Executive functions
Problem-solving
Inhibition
Working memory
Cognitive flexibility
Elementary school age

ABSTRACT

Previous research has shown that executive functions can contribute to successful problem-solving in preschool and elementary school children. However, most studies did not simultaneously assess the role of different specific aspects of executive functions. Therefore, the aim of our study was to investigate the individual contribution of inhibition, working memory, and cognitive flexibility to science problem-solving performance in elementary school children. A total of 478 children from first and second grades (M_{age} = 7.44 years) participated in our study. They performed a Go/No-go task (inhibition), a Corsi blocks backward task (working memory), a flexible item selection task (cognitive flexibility), and three science problem-solving tasks, including two gear turning tasks and one stabilization task. Structural equation modeling showed that working memory and cognitive flexibility individually contributed to problem-solving performance, whereas inhibition did not. We conclude that maintaining task requirements and dynamic object relations (working memory) and switching between different problem-solving phases (cognitive flexibility) are essential components of successful science problem-solving

* Corresponding author.
E-mail address: jonas.schaefer@rptu.de (J. Schäfer).

https://doi.org/10.1016/j.jecp.2024.105962

in elementary school children. Inhibitory processes may be more relevant in tasks involving a higher degree of interference at the task or response level.

Introduction

Executive functions are a set of higher-order control functions, including the three core functions inhibition, working memory, and cognitive flexibility (Diamond & Lee, 2011; Miyake et al., 2000). Their development is of major importance for cognitive and social development (Brocki & Bohlin, 2004; Moriguchi, 2014; Wiebe & Karbach, 2017), and previous research has highlighted their contribution to different facets of self-regulated and goal-oriented behavior (Diamond, 2013). Moreover, executive functions enable children to successfully use learning opportunities, as evidenced by the findings that executive functions predict reading and math abilities (Diamond, 2013; Gathercole, Pickering, Knight, & Stegmann, 2004b; Titz & Karbach, 2014; Johann, Könen, & Karbach, 2020). Previous research suggests that specific aspects of executive functions predict science achievement as well, although there is less evidence for the role of executive functions with respect to children's scientific literacy (Anthony & Ogg, 2019; Bauer & Booth, 2019).

Problem-solving is considered another higher-order cognitive skill that contributes to academic performance (Funke et al., 2018; Greiff et al., 2014). It refers to the cognitive processes involved in reaching a desired state that differs from a given state (Funke et al., 2018; Wang & Chiew, 2010). The problem-solving process is considered as a number of subsequent cognitive operations, including problem representation, planning (strategy formation), plan execution (strategy application), and outcome evaluation (Polya, 1957; Zelazo et al., 1997). These phases may be interdependent and recursive, for instance, because evaluating an outcome might bring insights that lead to a new problem representation and strategy formation (Dixon & Boncoddo, 2009).

Previous studies indicate that some aspects of executive functions are positively associated with problem-solving in early and middle childhood (Lee et al., 2009; Marulis & Nelson, 2021; McClelland & Cameron, 2019; Swanson & Beebe-Frankenberger, 2004; Viterbori et al., 2017). In particular, working memory was shown to be associated with problem-solving accuracy and the learning of problem-solving principles in preschool and elementary school children (Ropovik, 2014; Swanson & Beebe-Frankenberger, 2004; Swanson & Fung, 2016). Findings on the role of inhibition and cognitive flexibility in children's problem-solving are heterogeneous (Cañas et al., 2003; Robinson & Dubé, 2013; Viterbori et al., 2017). Nevertheless, problem-solving requires inhibition of task-irrelevant information and keeping track of problem representations, strategies, and outcomes, and it involves flexible attention shifts according to dynamically changing problem states (Zelazo et al., 1997). Given that these cognitive abilities are core aspects of executive functions, we hypothesize that executive functions are associated with problem-solving in children.

One domain of problem-solving that has received particular attention is science problem-solving (Taconis et al., 2001), which is closely related to science, technology, engineering, and mathematics (STEM) learning (Astuti et al., 2021). Children's problem-solving in STEM can entail manipulating or constructing object configurations (e.g., including gears and building blocks) that children are intrinsically motivated to engage with (Reuter & Leuchter, 2022; Weber & Leuchter, 2020). Therefore, science problem-solving tasks are frequently used in experimental studies with younger age groups such as preschoolers and children in the first grades of elementary school (Reuter & Leuchter, 2022; Schäfer et al., 2024a). However, there still is little evidence on the cognitive correlates of these science problem-solving tasks in elementary school children.

Executive functions

The three core executive functions (inhibition, working memory, and cognitive flexibility) develop significantly across childhood and well into adolescence (Wiebe & Karbach, 2017). They are excellent predictors for learning outcomes and academic performance (e.g., Johann & Karbach, 2021), and in turn they are impaired in many developmental and learning disorders (e.g., Barkley, 1997; Brandenburg et al., 2015).

Inhibition refers to the ability to suppress prepotent action tendencies and to control interference (Johnstone et al., 2007). It can support learning processes by allowing children to control interference from task-irrelevant stimuli or by helping them to take turns in the classroom. *Working memory* refers to the ability to maintain and manipulate information within the scope of seconds to a few minutes after stimulus presentation (Gathercole et al., 2004a). Therefore, it is involved in most mentally demanding situations (Ellis et al., 2020), for instance, during mental calculation involving large numbers and while following complex instructions. *Cognitive flexibility* refers to the ability to flexibly switch the attention between different task demands or activities (Titz & Karbach, 2014), for instance, when verbal and visual information is presented at the same time.

Executive functions in problem-solving

Problem-solving is considered an ability that is based on task-specific knowledge as well as domain-general cognitive abilities (Schäfer et al., 2024a; Funke et al., 2018). Many problems can be solved by retrieving task-specific knowledge and dealing with the situation routinely (Rahman, 2019). However, if task-specific knowledge is limited or missing, the problem-solving process may be less routine and more adaptive (Funke et al., 2018). Consequently, individuals lacking task-specific knowledge may increasingly rely on domain-general abilities such as reasoning and executive functions (Bühner et al., 2008; Greiff et al., 2016). Moreover, domain-general abilities enable the application of task-specific knowledge in sixth graders' problem-solving (Greiff et al., 2016).

Yet, evidence for the contribution of different aspects of executive control functions to elementary school children's problem-solving performance is scarce. Greiff et al. (2016) found that working memory had a lower impact on knowledge application than fluid reasoning in sixth-grade students, but the study did not assess other aspects of executive control such as inhibition and cognitive flexibility. Still, previous evidence suggests that executive functions may be of major importance for problem-solving in childhood because they are essential for inhibiting non-goal-directed behavior, maintaining and processing relevant information (Zelazo et al., 1997), and changing perspectives and strategies (Jonassen, 2000). Previous research has explored the differential contribution of executive functions to problem-solving to some extent but mainly in the mathematical problem domain. For instance, Viterbori et al. (2017) found that mathematical word problem-solving accuracy in 8-year-olds was predicted by working memory but not by inhibition and cognitive flexibility. Regarding the problem-solving phases, they found that inhibition and cognitive flexibility exclusively contributed to the planning phase and that working memory exclusively contributed to the execution phase. The phases of problem representation and outcome evaluation were not predicted by any executive function. Swanson and Beebe-Frankenberger (2004) found that working memory was the strongest predictor of word problem-solving performance in first to third graders. In their study, working memory explained unique variance in problem-solving performance beyond measures of inhibition, verbal, and calculation abilities. Inhibition did not positively contribute to problem-solving performance, and cognitive flexibility was not assessed (Swanson & Beebe-Frankenberger, 2004).

Exploring the relationship among problem-solving, computational thinking (CT), and executive functions provides additional information on the role of executive functions. *Computational thinking* refers to the specific cognitive skill set involved in processing complex problems algorithmically (Román-González et al., 2017; Yadav et al., 2017). Research suggests that CT and problem-solving abilities are closely related given that both involve complex reasoning, decomposition, and evaluation processes (Arfé et al., 2020). Current research has provided evidence that CT interventions (i.e., child-appropriate coding activities) may positively affect inhibition and working memory in children aged 4 to 16 years (Montuori et al., 2024). In particular, integrating coding activities in STEM lessons has shown to foster executive functions in first graders (Arfé et al., 2020). These findings may reflect that executive functions enable children to cognitively regulate the way they process problems, which leads to a more algorithmic procedure (Robertson et al., 2020). However, CT typically refers to the processing of problems in the technology context, whereas problem-solving is considered a more general term representing a broader cognitive framework (Greiff et al., 2014). Although CT abilities might be applied to real-world problems (Rabiee & Tjoa, 2017), problem-solving does not necessarily rely on algorithmic processes but may, for example, include the use of heuristics and a lack of routine knowledge in the problem-solver (Funke, 2012; Tönnsen, 2021).

Previous research has shown that executive functions are positively associated with children's science achievement both cross-sectionally (Bauer & Booth, 2019) and longitudinally (Anthony & Ogg, 2019). Moreover, executive functions may enable essential problem-solving processes, such as evaluating evidence by testing hypotheses and matching own predictions with actual outcomes, in preschoolers (Gropen et al., 2011). However, these studies assessed science achievement via question-based tasks but not via action-based construction tasks. Moreover, they either focused on preschoolers only (Bauer & Booth, 2019; Gropen et al., 2011) or did not assess inhibition (Anthony & Ogg, 2019). Even though problem-solving abilities are considered an essential science achievement (Greiff et al., 2014), and executive functions and problem-solving abilities develop across childhood (Brocki & Bohlin, 2004; Injoque-Ricle et al., 2014), there is still very little evidence on the association of executive functions and science problem-solving in elementary school age.

Aims and predictions

Consequently, the aims of our study were to investigate the contribution of inhibition, working memory, and cognitive flexibility to problem-solving performance in the science domain at early elementary school age. We drew the following predictions:

1. Inhibition may have an impact on problem-solving performance because it might support the suppression of task-irrelevant information or inappropriate rules or strategies.
2. Working memory may support problem-solving by facilitating to keep track of task requirements and to remember previously applied strategies and their outcomes as well as to recognize the logic of dynamic interdependencies between task-relevant objects.
3. Cognitive flexibility may contribute to problem-solving by supporting attentional switches between different rules and strategies as well as between different phases of the problem-solving process (i.e., problem representation, planning, execution, and evaluation). For instance, cognitive flexibility could enable the consideration of alternative problem-solving strategies if previous strategies were unsuccessful. Thus, we expected that all the core executive functions would explain unique variance in science problem-solving performance.

Method

Participants

A total of 478 first and second graders aged 6 to 8 years (M = 7.44 years, SD = 0.66; 228 female) participated in the study voluntarily and with written informed consent of their caregivers. The study was approved by the local ethics committee in the Department of Psychology at the University of Kaiserslautern–Landau. Data of 2 participants were excluded because they aborted participation prematurely because they did not feel well. Some children (3,97%) took extremely long to complete tasks and therefore were not able to finish the sessions in time (1 participant did not complete the Go/No-go task, 8 participants did not complete the Corsi blocks backward task, 3 participants did not complete the flexible item selection task (FIST), 8 participants did not complete the carousel task, and 1 participant did not complete the propeller task). This resulted in 459 complete data sets.

Procedure

The experiments were carried out in quiet rooms of the elementary schools in groups of 10 to 22 participants. All participants had a tablet on the table in front of them and worked individually on the tasks. Four to six trained experimenters and a teacher were present throughout the study. Task instructions were provided over headphones through app-integrated, professionally recorded audio tracks. At the beginning of the first session, the experimenters welcomed the participants and explained to them that they should address the experimenters or their teacher if they had difficulties with the tasks or wanted to stop participating. Each child was then assigned to an experimenter, who was available to answer the child's questions but did not give any hints beyond the task instructions provided in the audio tracks. Moreover, the experimenters explained to the participants that the study included solving games and puzzles on the tablet that involved gears, building blocks, and fairy tale characters. Subsequently, participants performed 13 tasks in two sessions (40 min each) separated by a 15-min break. The executive function tasks Go/No-go task (inhibition), Corsi blocks backward task (working memory), and FIST (cognitive flexibility) were performed in the first session. These tasks have been successfully applied in elementary school children in previous studies and were age appropriate in terms of difficulty, length, and complexity (Gathercole, Pickering, Ambridge, & Wearing, 2004a; Jacques & Zelazo, 2001; Johann & Karbach, 2018; St. John, Finch, & Tarullo, 2019). Given that our problem-solving tasks were action-based and did not tap the ability to resist to distractor interference, we chose a task assessing prepotent response inhibition (Go/No-go task) over a task measuring interference control (e.g., Eriksen flanker task; Eriksen & Eriksen, 1974). Because a child-friendly task presentation has been shown to foster children's performance in executive function tasks (Johann & Karbach, 2018), the executive function tasks were embedded in a cover story with fairy-tale-like characters. The three problem-solving tasks carousel, propeller, and stabilization were performed consecutively in the second session. In addition, tasks measuring mental rotation, fluid reasoning, and short-term memory (in the first session) as well as scientific concept understanding, vocabulary, and problem-solving transfer (in the second session) were administered, but they were not part of this study.

Materials

All tasks were administered digitally on a 10.1-inch-sized tablet with a 2560 × 1600-pixel resolution (Samsung Galaxy Note 10.1, Android Version 5.1.1). Audio tracks were implemented in the program to provide task instructions and explain the cover story to the participants. The tasks started after the instructions had been presented completely. The program was developed in Unity (Version 2020.3.17f1) and built via Android SDK (compression method LZ4). The problem-solving tasks were processed by single-point swipe control in a two-dimensional response space (see Schäfer et al., 2024b, for a more detailed specification), and the executive function tasks were processed by tapping selected responses.

Tasks

Inhibition: Go/No-go task (St. John et al., 2019)

Images of four different animals (cow, chicken, donkey, and pig) were presented one after the other for 500 ms each (see Fig. 1). Participants pressed a blue button below the animal when a chicken, donkey, or pig was presented (Go trials). However, if a cow was shown, they should withhold from pressing the button (No-go trials). In total, 60 Go trials and 20 No-go trials were presented in randomized order. After stimulus onset, participants had a time interval of 1000 ms to provide a response. After this response interval or after pressing the button (whichever came first), there was an interstimulus interval of 1000 ms followed by the presentation of the subsequent stimulus. Dependent variables were the normalized accuracy rate across Go and No-go trials and the inhibition score (equally weighted sum of the normalized hit rate and the normalized correct rejection rate).

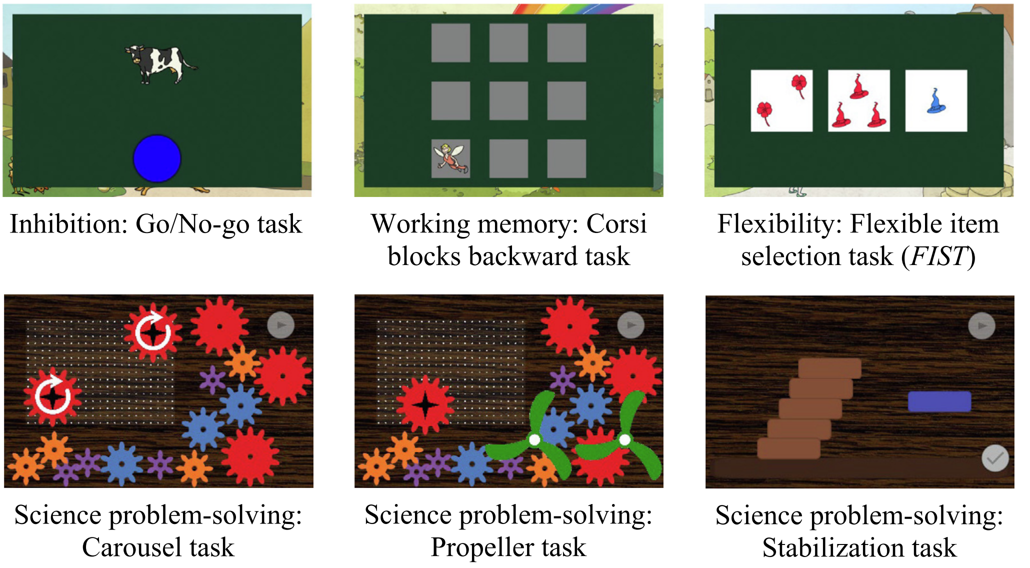

Fig. 1. Screenshots of the tasks measuring executive functions (upper panel) and science problem-solving (lower panel).

Working memory: Corsi blocks backward task (Gathercole et al., 2004a)

A matrix of 3 × 3 squares was presented to the participants. In each trial, a fairy appeared on some of the squares subsequently (see Fig. 1). This sequence needed to be recalled by the participants in reverse order. A sequence consisted of one marked square at the beginning of the task and could successively extend up to eight squares. Each square was presented at most once per sequence. For each sequence length, six sequences (i.e., trials) were presented. When fewer than four of six trials of one sequence length were recalled correctly, the task was aborted. Dependent variables were the product score (sequence length × number of correctly recalled trials) and the accuracy rate (correctly recalled squares in relation to all presented squares).

Cognitive flexibility: Flexible item selection task (Jacques & Zelazo, 2001)

The FIST consisted of a one-commonality phase (18 trials) and a consequent two-commonalities phase (9 trials). In the one-commonality phase, two pictures were presented that had in common a visual property in one of the three categories color (blue, red, yellow), number (one, two, three), or shape (hat, flower, fairy) (see Fig. 1). This common property was communicated to the participant (e.g., "These pictures are both blue"). Subsequently, a third picture was presented. Participants needed to pick the one of both initially presented pictures that shared a new common property with the third picture. In the two-commonalities phase, three pictures were presented simultaneously. Participants needed to choose the picture that shared one categorical property with each of the other two pictures. In both phases, each of the categories served equally often as the target category. Dependent variables were the rates of correct responses per task phase, that is, one-commonality accuracy rate and two-commonalities accuracy rate.

Problem-solving: Carousel task (Reuter & Leuchter, 2022)

In the carousel task, participants were to connect two gears that were fixed on a board by means of adding new gears in order to make both fixed gears turn in the same direction (see Fig. 1). To do this, the participants were provided with a sufficient number of movable gears of different sizes and up to 3 min of processing time. The dependent variable was computed based on the following scoring system (4 points: the fixed gears turned in the same direction; 3 points: the fixed gears were connected but did not turn in the same direction; 2 points: both fixed gears were connected to at least one other gear, but they were not connected to each other; 1 point: one fixed gear was connected to at least one gear, but the other was not; 0 points: neither fixed gear was connected to any gear).

Problem-solving: Propeller task (Schäfer et al., 2024a)

The propeller task consisted of fixing two propellers on movable gears so that one propeller turned as fast as possible and the other propeller turned as slowly as possible (see Fig. 1). Both propellers should be driven by a fixed starting gear without touching each other while turning. To do this, participants were provided with a sufficient number of movable gears of different sizes, two propellers of the same size, and up to 3 min of processing time. The dependent variable was defined as the sum of a turning speed score (3 points: one propeller was mounted on a largest sized gear and one propeller was mounted on a smallest sized gear; 2 points: propellers were mounted on different sized gears but not on largest and smallest; 1 point: both propellers were mounted on same sized gears; 0 points: one or both propellers were not attached to gears) and a contact score (2 points: propellers did not touch each other; 1 point: propellers touched each other; 0 points: starting gear did not drive both propellers).

Problem-solving: Stabilization task (Weber et al., 2020)

The stabilization task consisted of stabilizing unstable block constructions by adding another block (see Fig. 1). Participants had up to 3 attempts per item. The dependent variable was calculated as a budget of 24 points, from which 1 point per unsuccessful attempt was subtracted.

Statistical analyses

We first computed correlations (Spearman's rho) to test the association between all measures of executive functions and science problem-solving. The main analysis was based on a structural equation model with maximum likelihood estimation. We conducted a confirmatory factor analysis with the latent factors inhibition, working memory, cognitive flexibility, and science problem-solving. Accuracy rate and inhibition score served as indicators for the factor inhibition, accuracy rate and product score served as indicators for the factor working memory, and one-commonality accuracy rate and two-commonalities accuracy rate served as indicators for the factor cognitive flexibility. The factor science problem-solving was represented by the performance scores on the carousel, propeller, and stabilization tasks. To investigate the association of executive functions with science problem-solving abilities, we included a regression path from each executive function factor to the science problem-solving factor. According to Beauducel and Wittmann (2005), model fit was assessed by the χ^2 test, the comparative fit index (CFI), the Tucker–Lewis index (TLI), the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR). The initial loading of a latent factor was fixed to 1. All analyses were performed in R Version 4.3.0 (R Core Team, 2023).

Results

Descriptives and correlations of all dependent measures are shown in Table 1. Most measures significantly correlated. The measures selected to represent a common latent factor showed the expected substantial associations.

We then fitted a confirmatory factor model with one latent factor for each executive function and one latent factor for science problem-solving. The fit of the model to the observed data was excellent (see Fig. 2). Given that there is also evidence for a unitary structure of executive functions in middle childhood (Laureys et al., 2022), we also fitted a confirmatory factor model with a single latent factor representing all three executive functions. However, this model exhibited a poor model fit, $\chi^2(26) = 480.51$, $p = .000$, CFI = .64, TLI = .51, RMSEA = .19 (90% confidence interval [CI] = .18–.21), SRMR = .15. A likelihood ratio test showed that the fit of this single-factor model was significantly

Table 1
Descriptives and correlations between the executive functions and problem-solving variables

Variable	*M*	*SD*	Min	Max	1	2	3	4	5	6	7	8
1. Inhibition: Accuracy rate	0.79	0.13	0	1								
2. Inhibition: Inhibition score	1.50	0.22	0	2	.89**							
3. Working memory: Product score	44.78	37.51	0	216	.32**	.25**						
4. Working memory: Accuracy rate	0.65	0.16	0	1	.31**	.25**	.82**					
5. Cognitive flexibility: One-commonality accuracy rate	0.79	0.20	0	1	.30**	.27**	.41**	.36**				
6. Cognitive flexibility: Two-commonalities accuracy rate	0.51	0.22	0	1	.09	.08	.20**	.21**	.42**			
7. Science problem-solving: Carousel task performance score	2.85	1.08	0	4	.12**	.10*	.20**	.17**	.18**	.07		
8. Science problem-solving: Propeller task performance score	2.23	1.54	0	5	.18**	.14**	.29**	.25**	.33**	.15**	.26**	
9. Science problem-solving: Stabilization task performance score	11.16	3.30	0	24	.22**	.17**	.29**	.24**	.30**	.20**	.23**	.25**

Note. N = 476.
* $p < .05$.
** $p < .01$.

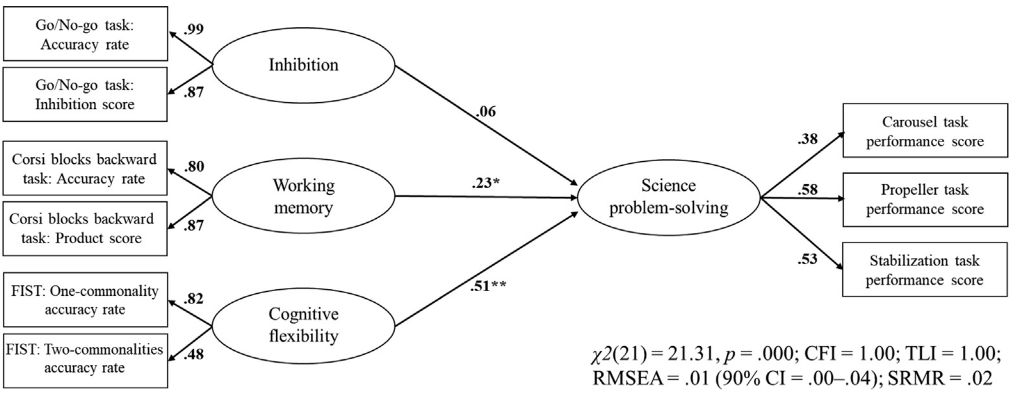

Fig. 2. Multiple regression model of executive functions and science problem-solving. The rectangles represent manifest variables, and the ovals represent latent variables. Single-headed arrows from latent variables to latent variables indicate regressions, and single-headed arrows from latent variables to manifest variables indicate factor loadings. CFI, comparative fit index; TLI, Tucker–Lewis index; RMSEA, root mean square error of approximation; CI, confidence interval; SRMR, standardized root mean square residual. *$p < .05$; **$p < .01$.

worse than the fit of the model with three latent executive function factors, $\Delta\chi^2(\Delta df = 5) = 459.20$, $p(\Delta\chi^2) < .001$.

The model showed that working memory and cognitive flexibility were significant predictors of science problem-solving and inhibition was not. Cognitive flexibility was a stronger predictor of science problem-solving than working memory, $p(\Delta\chi^2) < .001$, indicating that the contribution of cognitive flexibility to science problem-solving performance was significantly stronger than the contribution of working memory.

Discussion

Our research aim was to test whether the executive functions inhibition, working memory, and cognitive flexibility explained variance in problem-solving performance in elementary school children. The results of our study showed that most executive function measures were correlated with problem-solving performance on the manifest level. A latent confirmatory factor analysis showed that whereas inhibition was not significantly associated with problem-solving performance, working memory and cognitive flexibility were significantly related. Cognitive flexibility explained more variance than working memory.

The fact that working memory was significantly associated with problem-solving performance confirmed our expectations and is in line with previous research (Greiff et al., 2016; Viterbori et al., 2017). This finding suggests that working memory enables elementary school children to keep track of task requirements, previously applied strategies, and dynamic spatial interdependencies between different task-relevant objects. These object interdependencies are based on rule-based principles of turning direction and turning speed of connected gears and of ways to stabilize building block constructions (Schäfer et al., 2024a). Previous research further suggests that—besides solution accuracy—working memory is only associated with performance in the execution phase (Viterbori et al., 2017). To investigate which particular problem-solving phases are associated with working memory in our tasks, number and timing of gear turnings and building block replacements over the course of task processing could be analyzed in future studies.

Furthermore, our analyses showed that cognitive flexibility was not only significantly related but also most strongly associated with problem-solving performance. This finding suggests that cognitive flexibility enables elementary school children to efficiently switch between demands of different problem-solving phases. Furthermore, it is in line with studies indicating that cognitive flexibility can support problem-solving performance (Cañas et al., 2003; Ionescu, 2012), but it is contrary to findings of Viterbori et al. (2017), who did not find effects of cognitive flexibility on mathematical word problem-solving performance. However, our science problem-solving tasks required considerably more cognitive flexibility than mathematical word problem-solving tasks for three reasons. First, our tasks required switches between cognitive (problem representation and planning) and motor (execute planned actions) requirements, whereas motor requirements are typically minimal to absent in mathematical problems. Second, in our tasks the problem state was more dynamic (each action changed the current problem state without necessarily getting closer to the target state) than in mathematical word problem-solving tasks. Third, the problem-solving process was highly iterative (in each task participants had various attempts to solve the problem, which is typically not the case in mathematical problem-solving tasks).

Although inhibition can also promote problem-solving performance by suppressing task-irrelevant information and misleading strategies (Robinson & Dubé, 2013), it might not be completely surprising that problem-solving performance was not associated with inhibition in our study. In the problem-solving tasks we applied, all stimuli were task-relevant because all gears and building blocks could potentially be used to generate correct solutions. Thus, our finding is in line with previous research suggesting that inhibition does not affect problem-solving performance when no interference from task-irrelevant stimuli needs to be resolved (Lee et al., 2009). Moreover, this is compatible with the finding of Viterbori et al. (2017) that inhibition affects the problem-solving phase of planning but not the solution accuracy. To verify this interpretation, not only the inhibition of prepotent responses (as measured by the Go/No-go task) but also the resistance to distractor interference should be assessed in future studies (Rey-Mermet et al., 2018).

Taken together, our results suggest that executive functions, particularly cognitive flexibility and working memory, are closely related to science problem-solving performance in elementary school children. However, the individual association strength of the executive core functions varied considerably: Cognitive flexibility was more strongly associated with problem-solving performance than working memory, and inhibition did not play a significant role. In fact, this may be because we included both inhibition and working memory in our study. The finding that inhibition was not significantly associated with problem-solving performance confirms the findings from Swanson and Beebe-

Frankenberger (2004), indicating that working memory may overshadow the contribution of inhibition. An explanation for this finding is that working memory enables selective attention and consequently reduces the demand to filter distracting information (Marshall & Bays, 2013). Moreover, our results are partly in line with previous research suggesting a positive association of executive functions and CT (Arfé et al., 2020; Montuori et al., 2024). Accordingly, the findings may suggest a partial confirmation of a construct overlap between problem-solving and CT abilities (Román-González et al., 2017).

The positive associations of working memory and cognitive flexibility with problem-solving performance are in line with previous studies (Bühner et al., 2008; Greiff et al., 2016; Viterbori et al., 2017). Yet, it is noteworthy that cognitive flexibility was more strongly associated with problem-solving performance than working memory even though working memory was most strongly associated with problem-solving performance in previous studies (Bühner et al., 2008; Swanson & Beebe-Frankenberger, 2004; Viterbori et al., 2017). However, some of these studies did not include measures of cognitive flexibility (Bühner et al., 2008; Swanson & Beebe-Frankenberger, 2004). We conclude that children's cognitive flexibility is particularly relevant in hands-on science problem-solving due to dynamically changing problem states and problem-solving phases that require multiple attention shifts. Consequently, a replication of these findings in other problem-solving domains would be helpful in order to generalize the current findings.

Given that different aspects of executive functions come into play at different phases of problem-solving (cf. Arfé et al., 2020; Viterbori et al., 2017), it may also be expected that their individual contribution to problem-solving varies throughout the problem-solving process. Consequently, future research should give the contribution of working memory and cognitive flexibility to the different phases in science problem-solving closer consideration. Moreover, given that problem-solving strategies change across preschool and elementary school age (Injoque-Ricle et al., 2014), age-related changes in the individual contribution of distinct executive functions to problem-solving performance should be investigated in longitudinal studies.

Our study has the methodological limitation that the latent factors of the executive functions were represented by two variables derived from the same task. Although these variables represent different aspects of task performance, this may have led to an overestimation of the factor loadings.

Conclusion

Our study showed that the executive functions working memory and cognitive flexibility provide a significant contribution to science problem-solving performance in 6- to 8-year-old elementary school children. Consequently, our findings indicate that science problem-solving, as a central STEM competence, is closely related to executive functions. This suggests that cognitive control functions are central skills in children's STEM performance.

CRediT authorship contribution statement

Jonas Schäfer: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization, Project administration. **Timo Reuter:** Writing – review & editing, Investigation. **Miriam Leuchter:** Writing – review & editing, Resources, Project administration, Funding acquisition. **Julia Karbach:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Data availability

Data will be made available on request.

References

Anthony, C. J., & Ogg, J. (2019). Executive function, learning-related behaviors, and science growth from kindergarten to fourth grade. *Journal of Educational Psychology, 112*(8), 1563–1581. https://doi.org/10.1037/edu0000447.

Arfé, B., Vardanega, T., & Ronconi, L. (2020). The effects of coding on children's planning and inhibition skills. *Computers & Education, 148*, 103807. https://doi.org/10.1016/j.compedu.2020.103807.

Astuti, N. H., Rusilowati, A., & Subali, B. (2021). STEM-based learning analysis to improve students' problem solving abilities in science subject: A literature review. *Journal of Innovative Science Education, 10*(1), 79–86. https://doi.org/10.15294/JISE.V9I2.38505.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*(1), 65–94. https://doi.org/10.1037/0033-2909.121.1.65.

Bauer, J.-R., & Booth, A. E. (2019). Exploring potential cognitive foundations of scientific literacy in preschoolers: Causal reasoning and executive function. *Early Childhood Research Quarterly, 46*, 275–284. https://doi.org/10.1016/j.ecresq.2018.09.007.

Beauducel, A., & Wittmann, W. W. (2005). Simulation study on fit indexes in CFA based on data with slightly distorted simple structure. *Structural Equation Modeling: A Multidisciplinary Journal, 12*(1), 41–75. https://doi.org/10.1207/s15328007sem1201_3.

Brandenburg, J., Klesczewski, J., Fischbach, A., Schuchardt, K., Büttner, G., & Hasselhorn, M. (2015). Working memory in children with learning disabilities in reading versus spelling: Searching for overlapping and specific cognitive factors. *Journal of Learning Disabilities, 48*(6), 622–634. https://doi.org/10.1177/0022219414521.

Brocki, K. C., & Bohlin, G. (2004). Executive functions in children aged 6 to 13: A dimensional and developmental study. *Developmental Neuropsychology, 26*(2), 571–593. https://doi.org/10.1207/s15326942dn2602_3.

Bühner, M., Kröner, S., & Ziegler, M. (2008). Working memory, visual–spatial-intelligence and their relationship to problem-solving. *Intelligence, 36*(6), 672–680. https://doi.org/10.1016/j.intell.2008.03.008.

Cañas, J., Quesada, J., Antolí, A., & Fajardo, I. (2003). Cognitive flexibility and adaptability to environmental changes in dynamic complex problem-solving tasks. *Ergonomics, 46*(5), 482–501. https://doi.org/10.1080/0014013031000061640.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750.

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science, 333*(6045), 959–964. https://doi.org/10.1126/science.1204529.

Dixon, J. A., & Boncoddo, R. (2009). Strategies and problem representations: Implications for models of changing cognitive structure—Commentary on "Learning new problem-solving strategies leads to changes in problem representation" by M. W. Alibali, K. M. Ockuly, and A. D. Fischer. *Cognitive Development, 24*(2), 102–105. https://doi.org/10.1016/j.cogdev.2009.03.002.

Ellis, D. M., Ball, B. H., Kimpton, N., & Brewer, G. A. (2020). The role of working memory capacity in analytic and multiply-constrained problem-solving in demanding situations. *Quarterly Journal of Experimental Psychology, 73*(6), 920–928. https://doi.org/10.1177/1747021820909703.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*, 143–149. https://doi.org/10.3758/BF03203267.

Funke, J. (2012). Complex problem solving. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 682–685). Springer. https://doi.org/10.1007/978-1-4419-1428-6_685.

Funke, J., Fischer, A., & Holt, D. V. (2018). Competencies for complexity: Problem solving in the twenty-first century. In E. Care, P. Griffin, & M. Wilson (Eds.), *Educational assessment in an information age: Assessment and teaching of 21st century skills* (pp. 41–53). Springer. https://doi.org/10.1007/978-3-319-65368-6_3.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004a). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*(2), 177–190. https://doi.org/10.1037/0012-1649.40.2.177.

Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004b). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology, 18*(1), 1–16. https://doi.org/10.1002/acp.934.

Greiff, S., Krkovic, K., & Hautamäki, J. (2016). The prediction of problem-solving assessed via microworlds. *European Journal of Psychological Assessment, 32*(4), 298–306. https://doi.org/10.1027/1015-5759/a000263.

Greiff, S., Wüstenberg, S., Csapó, B., Demetriou, A., Hautamäki, J., Graesser, A. C., & Martin, R. (2014). Domain-general problem solving skills and education in the 21st century. *Educational Research Review, 13*, 74–83. https://doi.org/10.1016/j.edurev.2014.10.002.

Gropen, J., Clark-Chiarelli, N., Hoisington, C., & Ehrlich, S. B. (2011). The importance of executive function in early science education. *Child Development Perspectives, 5*(4), 298–304. https://doi.org/10.1111/j.1750-8606.2011.00201.x.

Injoque-Ricle, I., Barreyro, J. P., Calero, A., & Burin, D. I. (2014). Tower of London: Planning development in children from 6 to 13 years of age. *Spanish Journal of Psychology, 17*, E77. https://doi.org/10.1017/sjp.2014.83.

Ionescu, T. (2012). Exploring the nature of cognitive flexibility. *New Ideas in Psychology, 30*(2), 190–200. https://doi.org/10.1016/j.newideapsych.2011.11.001.

Jacques, S., & Zelazo, P. D. (2001). The Flexible Item Selection Task (FIST): A measure of executive function in preschoolers. *Developmental Neuropsychology, 20*(3), 573–591. https://doi.org/10.1207/S15326942DN2003_2.

Johann, V. E., & Karbach, J. (2018). Validation of new online game-based executive function tasks for Children. *Journal of Experimental Child Psychology, 176*, 150–161. https://doi.org/10.1016/j.jecp.2018.07.009.

Johann, V. E., & Karbach, J. (2021). Educational application of cognitive training. In T. Strobach & J. Karbach (Eds.), *Cognitive Training* (pp. 333–350). Cham: Springer. https://doi.org/10.1007/978-3-030-39292-5_23.

Johann, V. E., Könen, T., & Karbach, J. (2020). The unique contribution of working memory, inhibition, cognitive flexibility, and intelligence to reading comprehension and reading speed. *Child Neuropsychology, 26*(3), 324–344. https://doi.org/10.1080/09297049.2019.1649381.

Johnstone, S. J., Dimoska, A., Smith, J. L., Barry, R. J., Pleffer, C. B., Chiswick, D., & Clarke, A. R. (2007). The development of stop-signal and go/nogo response inhibition in children aged 7–12 years: Performance and event-related potential indices. *International Journal of Psychophysiology, 63*(1), 25–38. https://doi.org/10.1016/j.ijpsycho.2006.07.001.
Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development, 48*, 63–85. https://doi.org/10.1007/BF02300500.
Laureys, F., de Waelle, S., Barendse, M. T., Lenoir, M., & Deconinck, F. J. (2022). The factor structure of executive function in childhood and adolescence. *Intelligence, 90*, 101600. https://doi.org/10.1016/j.intell.2021.101600.
Lee, K., Ng, E. L., & Ng, S. F. (2009). The contributions of working memory and executive functioning to problem representation and solution generation in algebraic word problems. *Journal of Educational Psychology, 101*(2), 373–387. https://doi.org/10.1037/a0013843.
Marshall, L., & Bays, P. M. (2013). Obligatory encoding of task-irrelevant features depletes working memory resources. *Journal of Vision, 13*(2), Article 21. https://doi.org/10.1167/13.2.21.
Marulis, L. M., & Nelson, L. J. (2021). Metacognitive processes and associations to executive function and motivation during a problem-solving task in 3–5 year olds. *Metacognition and Learning, 16*(1), 207–231. https://doi.org/10.1007/s11409-020-09244-6.
McClelland, M. M., & Cameron, C. E. (2019). Developing together: The role of executive function and motor skills in children's early academic lives. *Early Childhood Research Quarterly, 46*, 142–151. https://doi.org/10.1016/j.ecresq.2018.03.014.
Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734.
Montuori, C., Gambarota, F., Altoé, G., & Arfé, B. (2024). The cognitive effects of computational thinking: A systematic review and meta-analytic study. *Computers & Education, 210*, 104961. https://doi.org/10.1016/j.compedu.2023.104961.
Moriguchi, Y. (2014). The early development of executive function and its relation to social interaction: A brief review. *Frontiers in Psychology, 5*, 388. https://doi.org/10.3389/fpsyg.2014.00388.
Polya, G. (1957). *How to solve it: A new aspect of mathematical method* (2nd ed.). Princeton University Press.
R Core Team. (2023). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org.
Rabiee, M., & Tjoa, A. M. (2017). From abstraction to implementation: Can computational thinking improve complex real-world problem solving? A computational thinking-based approach to the SDGs. In J. Choudrie, M. S. Islam, F. Wahid, J. M. Bass, & J. E. Priyatma (Eds.), *IFIP advances in information and communication technology, Vol. 504: Information and communication technologies for development* (pp. 104–116). Springer. https://doi.org/10.1007/978-3-319-59111-7_10.
Rahman, M. M. (2019). 21st century skill "problem solving": Defining the concept. *Asian Journal of Interdisciplinary Research, 2*(1), 64–74. https://doi.org/10.34256/ajir1917.
Reuter, T., & Leuchter, M. (2022). Examining Kindergarten Children's testing and optimising in the context of a gear engineering task. *European Journal of STEM Education, 7*(1), 04. https://doi.org/10.20897/ejsteme/11827.
Rey-Mermet, A., Gade, M., & Oberauer, K. (2018). Should we stop thinking about inhibition? Searching for individual and age differences in inhibition ability. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(4), 501–526. https://doi.org/10.1037/xlm0000450.
Robertson, J., Gray, S., Toye, M., & Booth, J. (2020). The relationship between executive functions and computational thinking. *International Journal of Computer Science Education in Schools, 3*(4), 35–49. https://doi.org/10.21585/ijcses.v3i4.76.
Robinson, K. M., & Dubé, A. K. (2013). Children's additive concepts: Promoting understanding and the role of inhibition. *Learning and Individual Differences, 23*, 101–107. https://doi.org/10.1016/j.lindif.2012.07.016.
Román-González, M., Pérez-González, J.-C., & Jiménez-Fernández, C. (2017). Which cognitive abilities underlie computational thinking? Criterion validity of the Computational Thinking Test. *Computers in Human Behavior, 72*, 678–691. https://doi.org/10.1016/j.chb.2016.08.047.
Ropovik, I. (2014). Do executive functions predict the ability to learn problem-solving principles? *Intelligence, 44*, 64–74. https://doi.org/10.1016/j.intell.2014.03.002.
Schäfer, J., Reuter, T., Karbach, J., & Leuchter, M. (2024a). Domain-specific knowledge and domain-general abilities in children's science problem-solving. *British Journal of Educational Psychology, 94*(2), 346–366. https://doi.org/10.1111/bjep.12649.
Schäfer, J., Reuter, T., Leuchter, M., & Karbach, J. (2024b). *Validation of new tablet-based problem-solving tasks in primary school students*. Manuscript submitted for publication.
St. John, A. M., Finch, K., & Tarullo, A. R. (2019). Socioeconomic status and neural processing of a go/no-go task in preschoolers: An assessment of the P3b. *Developmental Cognitive Neuroscience, 38*, 100677. https://doi.org/10.1016/j.dcn.2019.100677.
Swanson, H. L., & Beebe-Frankenberger, M. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology, 96*(3), 471–491. https://doi.org/10.1037/0022-0663.96.3.471.
Swanson, H. L., & Fung, W. (2016). Working memory components and problem-solving accuracy: Are there multiple pathways? *Journal of Educational Psychology, 108*(8), 1153–1177. https://doi.org/10.1037/edu0000116.
Taconis, R., Ferguson-Hessler, M. G. M., & Broekkamp, H. (2001). Teaching science problem solving: An overview of experimental work. *Journal of Research in Science Teaching, 38*(4), 442–468. https://doi.org/10.1002/tea.1013.
Titz, C., & Karbach, J. (2014). Working memory and executive functions: Effects of training on academic achievement. *Psychological Research, 78*, 852–868. https://doi.org/10.1007/s00426-013-0537-1.
Tönnsen, K.-C. (2021). The relevance of trial-and-error: Can trial-and-error be a sufficient learning method in technical problem-solving-contexts? *Techne Serien—Forskning i slöjdpedagogik och slöjdvetenskap, 28*(2), 303–312.
Viterbori, P., Traverso, L., & Usai, M. C. (2017). The role of executive function in arithmetic problem-solving processes: A study of third graders. *Journal of Cognition and Development, 18*(5), 595–616. https://doi.org/10.1080/15248372.2017.1392307.
Wang, Y., & Chiew, V. (2010). On the cognitive process of human problem solving. *Cognitive Systems Research, 11*(1), 81–92. https://doi.org/10.1016/j.cogsys.2008.08.003.

Weber, A. M., & Leuchter, M. (2020). Measuring preschool children's knowledge of the principle of static equilibrium in the context of building blocks: Validation of a test instrument. *The British Journal of Educational Psychology, 90*(1), 50–74. https://doi.org/10.1111/bjep.12304.

Weber, A. M., Reuter, T., & Leuchter, M. (2020). The impact of a construction play on 5- to 6-year-old Children's reasoning about stability. *Frontiers in Psychology, 11*, 1737. https://doi.org/10.3389/fpsyg.2020.01737.

Wiebe, S. A., & Karbach, J. (Eds.). (2017). *Executive Function: Development Across the Life Span* (1st ed.). New York:Routledge. https://doi.org/10.4324/9781315160719.

Yadav, A., Good, J., Voogt, J., & Fisser, P. (2017). Computational thinking as an emerging competence domain. In M. Mulder (Ed.), *Technical and vocational education and training: issues, concerns and prospects, Vol. 23: Competence-based vocational and professional education* (pp. 1051–1067). Springer. https://doi.org/10.1007/978-3-319-41713-4_49.

Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology, 1*(2), 198–226. https://doi.org/10.1037/1089-2680.1.2.198.

8 Appendix

A General Statement and Author Contributions (in German)

Eidesstattliche Erklärung

Hiermit erkläre ich, Jonas Schäfer, dass ich die Synopse der vorliegenden Dissertation eigenständig ohne die unzulässige Inanspruchnahme Dritter verfasst habe und keine anderen als die angegeben Hilfsmittel verwendet habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken habe ich als solche gekennzeichnet. Die vorliegende Dissertation ist weder im Ganzen noch in Teilen bei einer anderen Hochschule eingereicht worden. Für die drei im Rahmen dieser Dissertation verfassten Publikationen wurden folgende individuelle Beiträge von den einzelnen Autorinnen und Autoren (definiert nach dem CRediT-System) erbracht:

Artikel 1: Schäfer, J., Reuter, T., Leuchter, M., & Karbach, J. (2024). Validation of new tablet-based problem-solving tasks in primary school students. *PloS ONE*, *19*(8), e0309718. https://doi.org/10.1371/journal.pone.0309718

CRediT Role	Author 1: Jonas Schäfer	Author 2: Timo Reuter	Author 3: Miriam Leuchter	Author 4: Julia Karbach
1. Conceptualization	☒ equal	☒ equal	☒ equal	☒ equal
2. Data curation	☒ --	☐ --	☐ --	☐ --
3. Formal analysis	☒ --	☐ --	☐ --	☐ --
4. Funding acquisition	☐ --	☐ --	☒ equal	☒ equal
5. Investigation	☒ lead	☒ supporting	☐ --	☐ --
6. Methodology	☐ --	☐ --	☐ --	☐ --
7. Project administration	☐ --	☐ --	☐ --	☐ --
8. Resources	☐ --	☐ --	☒ equal	☒ equal
9. Software	☒ --	☐ --	☐ --	☐ --
10. Supervision	☐ --	☐ --	☐ --	☒ --
11. Validation	☒ --	☐ --	☐ --	☐ --
12. Visualization	☒ --	☐ --	☐ --	☐ --
13. Writing – original draft	☒ --	☐ --	☐ --	☐ --
14. Writing – review & editing	☒ lead	☒ supporting	☒ supporting	☒ lead

Artikel 2: Schäfer, J., Reuter, T., Karbach, J., & Leuchter, M. (2024). Domain-specific knowledge and domain-general abilities in children's science problem-solving. *British Journal of Educational Psychology*, *94*(2), 346-366. https://doi.org/10.1111/bjep.12649

CRediT Role	Author 1: Jonas Schäfer	Author 2: Timo Reuter	Author 3: Julia Karbach	Author 4: Miriam Leuchter
1. Conceptualization	☒ equal	☒ equal	☒ equal	☒ equal
2. Data curation	☒ --	☐ --	☐ --	☐ --
3. Formal analysis	☒ --	☐ --	☐ --	☐ --
4. Funding acquisition	☐ --	☐ --	☐ --	☒ --
5. Investigation	☒ lead	☒ supporting	☐ --	☐ --
6. Methodology	☒ equal	☒ equal	☐ --	☐ --
7. Project administration	☐ --	☐ --	☐ --	☒ --
8. Resources	☐ --	☐ --	☒ equal	☒ equal
9. Software	☒ --	☐ --	☐ --	☐ --
10. Supervision	☐ --	☐ --	☐ --	☒ --
11. Validation	☒ --	☐ --	☐ --	☐ --
12. Visualization	☒ --	☐ --	☐ --	☐ --
13. Writing – original draft	☒ --	☐ --	☐ --	☐ --
14. Writing – review & editing	☒ lead	☒ lead	☒ supporting	☒ lead

Artikel 3: Schäfer, J., Reuter, T., Leuchter, M., & Karbach, J. (2024). Executive functions and problem-solving—The contribution of inhibition, working memory, and cognitive flexibility to science problem-solving performance in elementary school students. *Journal of Experimental Child Psychology*, *244*, 105962. https://doi.org/10.1016/j.jecp.2024.105962

CRediT Role	Author 1: Jonas Schäfer	Author 2: Timo Reuter	Author 3: Miriam Leuchter	Author 4: Julia Karbach
1. Conceptualization	☒ equal	☐ --	☐ --	☒ equal
2. Data curation	☒ --	☐ --	☐ --	☐ --
3. Formal analysis	☒ lead	☐ --	☐ --	☒ supporting
4. Funding acquisition	☐ --	☐ --	☒ equal	☒ equal
5. Investigation	☒ lead	☒ supporting	☐ --	☐ --
6. Methodology	☒ equal	☐ --	☐ --	☒ equal
7. Project administration	☒ equal	☐ --	☒ equal	☒ equal
8. Resources	☐ --	☐ --	☒ --	☐ --
9. Software	☒ --	☐ --	☐ --	☐ --
10. Supervision	☐ --	☐ --	☐ --	☒ --
11. Validation	☒ --	☐ --	☐ --	☐ --
12. Visualization	☒ --	☐ --	☐ --	☐ --
13. Writing – original draft	☒ --	☐ --	☐ --	☐ --
14. Writing – review & editing	☒ lead	☒ supporting	☒ supporting	☒ lead

B Acknowledgements (in German)

Danksagung

Zuerst bedanke ich mich ganz herzlich bei Prof. Dr. Miriam Leuchter und Prof. Dr. Julia Karbach für die Betreuung meiner Promotion. Ihre zuverlässige Unterstützung hat mich nicht nur fachlich, sondern auch persönlich weiterentwickelt. Prof. Dr. Miriam Leuchter möchte ich ganz besonders für ihre verständnisvolle Art und ihre außergewöhnliche Kreativität danken. Prof. Dr. Julia Karbach bin ich besonders dankbar für ihre Wertschätzung und ihren Optimismus, die mich immer wieder motiviert haben.

Prof. Dr. Daniela Czernochowski möchte ich für die Bereitschaft danken, sich mit ihrer großen kognitionspsychologischen Expertise als Gutachterin bereitzuerklären.

Ich bedanke mich bei allen engagierten studentischen Hilfskräften für ihre Unterstützung bei den Datenerhebungen. Mein Dank gilt auch allen Kindern für die Studienteilnahme sowie den Lehrkräften und Erziehungsberechtigten für das entgegengebrachte Vertrauen.

Vielen Dank an meine Kolleginnen und Kollegen aus den Arbeitsgruppen Entwicklungspsychologie und pädagogische Psychologie und Grundschulpädagogische Forschung der RPTU für den Austausch und die gemeinsamen Erlebnisse. Sie haben die Promotion für mich zu einer unterhaltsamen Zeit gemacht, auf die ich gerne zurückblicken werde.

C Academic Curriculum Vitae (in German)

Persönliche Daten

Name Jonas Schäfer

Akademische Berufserfahrungen

seit 01/2021 **Wissenschaftlicher Mitarbeiter**

AG Entwicklungspsychologie und pädagogische Psychologie & Grundschulpädagogische Forschung, Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau (bis 12/2022 Universität Koblenz-Landau)

Schulische & Akademische Ausbildung

03/2020 **M.Sc. Kognitionswissenschaft**

Eberhard Karls Universität, Tübingen

07/2016 **B.A. Sprache und Kommunikation**

Philipps-Universität, Marburg

07/2011 **Allgemeine Hochschulreife**

Gymnasium Dr. Richter, Kelkheim

Publikationen (Peer-Reviewed)

Schäfer, J., Reuter, T., Leuchter, M., & Karbach, J. (2024). Validation of new tablet-based problem-solving tasks in primary school students. *PLoS ONE, 19*(8), e0309718. doi:10.1371/journal.pone.0309718

Schäfer, J., Reuter, T., Karbach, J. & Leuchter, M., (2024). Domain-specific knowledge and domain-general abilities in children's science problem-solving. *British Journal of Educational Psychology, 94*(2), 346–366. doi:10.1111/bjep.12649

Schäfer, J., Reuter, T., Leuchter, M., & Karbach, J. (2024). Executive functions and problem-solving – the contribution of inhibition, working memory, and cognitive flexibility to science problem-solving performance in elementary school students. *Journal of Experimental Child Psychology, 244,* 105962. doi:10.1016/j.jecp.2024.105962

Konferenzbeiträge

Schäfer, J., Reuter, T., Leuchter, M., & Karbach, J. (07/2024). The association of executive functions and science problem-solving in primary school age. *Talk at the EARLI SIG5&28 Conference 2024, Warsaw, Poland*.

Schäfer, J., Reuter, T., Karbach, J., & Leuchter, M. (06/2024). The impact of spatial skills on engineering in primary school children. *Talk at the JURE 2024 Conference, Sevilla, Spain*.

Reuter, T., **Schäfer, J.,** & Leuchter, M. (04/2024). Fostering preschoolers' problem-solving with gears through gestural and verbal modelling. *Talk at the AERA Annual Meeting 2024, Philadelphia, USA.*

Reuter, T., **Schäfer, J.,** & Leuchter, M. (03/2024). Veränderung der Vorstellung zur Drehrichtung von Zahnrädern durch gestisches und sprachliches Modellieren bei Vorschulkindern. *Talk at the GEBF Conference 2024, Potsdam, Germany.*

Schäfer, J., Reuter, T., Leuchter, M., & Karbach, J. (09/2023). Executive functions and problem-solving: The impact of inhibition, flexibility, and working memory on technical problem-solving in primary school age. *Poster presented at the ESCoP Conference 2023, Porto, Portugal.*

Schäfer, J., Reuter, T., Karbach, J., & Leuchter, M. (08/2023). Domain-general and domain-specific competencies in children's spatial problem-solving. *Talk at the EARLI Conference 2023, Thessaloniki, Greece.*

Schäfer, J., Reuter, T., Leuchter, M., & Karbach, J. (08/2022). Problem-solving in middle childhood: Validation of a tablet-based assessment tool. *Poster presented at the ESCoP Conference 2022, Lille, France.*

Schäfer, J., Reuter, T., Karbach, J., & Leuchter, M. (07/2022). Domain-general and domain-specific components in young children's problem-solving. *Talk at the EARLI SIG5 Conference 2022, Utrecht, Netherlands.*

Kenntnisse

EDV-Kenntnisse	Java (sehr gut)
	R (sehr gut)
	C# (gut)
	SPSS (gut)
	Matlab & Psychtoolbox (gut)
Sprachkenntnisse	Englisch (fließend)
	Portugiesisch (Grundkenntnisse)
	Französisch (Grundkenntnisse)

Landau, 25.09.2024