Cognitive Processes in Children's Strategy Use in Computational Estimation

Dissertation
for the Degree of
Doctor of Natural Sciences

Submitted to the Department of Psychology and Sports Sciences
of the Goethe University Frankfurt

by
Svenja Hammerstein, M.Sc.
from Bad Soden, Germany

Frankfurt a. M., 2020

(D30)

Submitted to the Department of Psychology and Sports Sciences of the Goethe University Frankfurt as Dissertation for the Degree of Doctor of Natural Sciences

Dean:

Prof. Dr. Sonja Rohrmann

Reviewers:

Prof. Dr. Gerhard Büttner

Prof. Dr. Marcus Hasselhorn

Date of Disputation: 11.11.2020

TABLE OF CONTENTS

GERMAN SUMMARY (ZUSAMMENFASSUNG)

In alltäglichen Situationen nutzen wir fortwährend Strategien, um kognitive Aufgaben zu bewältigen. So wenden wir beispielsweise Gedächtnisstrategien an, um uns die Inhalte einer Einkaufsliste zu merken, ohne einen Einkaufszettel schreiben zu müssen. Als ein anderes Beispiel nutzen wir Überschlagsrechenstrategien, um zu bestimmen, wie viel wir am Ende des Einkaufs ungefähr für die Lebensmittel bezahlen müssen. Überschlagsrechnen ist jedoch nicht nur ein wichtiger Bestandteil des alltäglichen Lebens, sondern nimmt auch elementare Funktionen im schulischen Kontext ein. So hängt die Fähigkeit von Kindern im Überschlagsrechnen nicht nur mit ihrer Intelligenz (Reys, Rybolt, Bestgen, & Wyatt, 1982) sowie Ergebnissen in arithmetischen Schulleistungstests (Siegler & Booth, 2004) zusammen, sondern hilft Kindern auch, ein besseres Verständnis von anderen mathematischen Konstrukten zu erlangen (Star, Rittle-Johnson, Lynch, & Perova, 2009). Um Überschlagsrechenergebnisse zu erhalten, können Kinder verschiedene Strategien anwenden, welche die genaue Rechnung vereinfachen. Studien konnten zeigen, dass insbesondere jüngere Kinder auf Rundungsstrategien zurückgreifen, um Überschlagsrechnungen zu absolvieren (LeFevre, Greenham, & Waheed, 1993; Lemaire, Lecacheur, & Farioli, 2000). Überdies fanden Studien, dass Personen verschiedene Rundungsstrategien nutzen und diese adaptiv an Aufgabenmerkmale anpassen (Lemaire & Lecacheur, 2002; Xu, Wells, LeFevre, & Imbo, 2014). Verschiedene theoretische Modelle (Lovett & Anderson, 1996; Rieskamp & Otto, 2006; Shrager & Siegler, 1998) gehen davon aus, dass die Strategiewahl bei der Bearbeitung kognitiver Aufgaben assoziative Prozesse beinhaltet. Relative Kosten und Nutzen verfügbarer Strategien werden hierbei aktiviert und diejenige Strategie mit dem größten assoziierten Nutzen sowie den geringsten assoziierten Kosten gewählt. So werden beispielsweise die Genauigkeit des Überschlagsergebnisses sowie die Komplexität der Strategieprozedur in die Strategiewahl einbezogen.

Trotz des wachsenden Interesses an Überschlagsrechenstrategien von Kindern in den letzten Jahren, ist das Wissen über die dahinter liegenden kognitiven Prozesse sowie einflussnehmende Faktoren begrenzt. Daher lag der Fokus der vorliegenden Arbeit auf ebenjenen kognitiven Prozessen während der Nutzung von Überschlagsstrategien. Hierzu wurden in vier Studien zwei zentrale Forschungsinhalte verfolgt: (1) Wie wirken sich

individuelle Unterschiede in exekutiven Funktionen auf Überschlagsrechenstrategien aus und (2) wie wirkt sich die Variation von verschiedenen Aufgabenmerkmalen auf Überschlagsstrategien aus, welche mit kognitiven Prozessen assoziiert sind?

In den letzten Jahrzehnten konnten zahlreiche Studien zeigen, dass exekutive Funktionen eine wichtige Rolle für schulische Leistungen einnehmen. Exekutive Funktionen sind höhergeordnete Prozesse, welche andere kognitive Prozesse während zielgerichteten Handelns steuern (Friedman & Miyake, 2017). Drei zentrale exekutiven Funktionen, auf welche in der Literatur zumeist Bezug genommen wird, sind Inhibition, Shifting und Arbeitsgedächtnis Updating (Diamond, 2013; Miyake et al., 2000). Inhibition beschreibt Prozesse der Hemmung automatischer oder dominanter Reaktionen sowie des Ausblendens irrelevanter Reize. Shifting bezieht sich auf die Fähigkeit, zwischen verschiedenen Aufgaben oder kognitiven Anforderungen zu wechseln. Updating beschreibt die Koordination sowie Manipulation von Arbeitsgedächtnisinhalten. Im arithmetischen Bereich scheint insbesondere die exekutive Funktion des Updatings einen zentralen Einfluss auszuüben (Friso-van den Bos, van der Ven, Kroesbergen, & van Luit, 2013; Lechuga, Pelegrina, Pelaez, Martin-Puga, & Justicia, 2016; St Clair-Thompson, & Gathercole, 2006; van der Ven, Kroesbergen, Boom, & Leseman, 2012). Trotz dieser Befunde wurde jedoch der Einfluss von Updating bislang in keiner Studie zum arithmetischen Strategiegebrauch untersucht. Daher lag der Fokus von Studie 1 auf der Rolle von Updating für den Rundungsstrategiegebrauch bei Grundschulkindern. Hierzu bearbeiteten 158 Drittklässler[1] sowie 150 Viertklässler Überschlagsrechenaufgaben. Sie wurden darin angeleitet, eine von zwei Rundungsstrategie zu nutzen: die Abrundungsstrategie (d.h. beide Summanden abzurunden) oder die Aufrundungsstrategie (d.h. beide Summanden aufzurunden). Die Überschlagsaufgaben setzten sich aus zwei Aufgabentypen zusammen: homogene Aufgaben (d.h. die Einerstellen der Summanden sind beide kleiner oder beide größer als fünf) sowie heterogenen Aufgaben (d.h. die Einerstelle eines Summanden ist kleiner als fünf und die Einerstelle des anderen Summanden ist größer als fünf). Ferner absolvierten die Kinder vier Updating Aufgaben. Die Ergebnisse zeigten, dass Kinder mit höheren Updating Leistungen die beiden

[1] In der vorliegenden Arbeit wurden Substantive mit unterschiedlichen Formen für die Geschlechter zur besseren Lesbarkeit vermieden. Wenn keine geschlechtsneutrale Formulierung passend war, wurde die grammatikalisch männliche Form gewählt, welche sich als generisches Maskulinum auf beide Geschlechter beziehen kann.

Rundungsstrategien flexibler anwandten. Überdies wählten diese Kinder häufiger die bessere der beiden Strategien und führten die Strategien schneller aus – insbesondere bei numerisch größeren Aufgaben, bei homogenen Problemen sowie bei der Anwendung der komplexeren Aufrundungsstrategie. Die Ergebnisse sprechen dafür, dass Updating eine wichtige Rolle bei der Bearbeitung von Überschlagsaufgaben einnimmt. Insbesondere bei der Aktivierung von assoziativen Prozessen scheint effizientes Updating die Konsolidierung von Langezeitgedächtnisinhalten und den Zugriff auf ebenjene zu vereinfachen. Überdies scheinen komplexere Aufgaben für Kinder mit besserer Updating Leistung mit geringeren relativen Kosten assoziiert zu sein.

Wie auch in der ersten Studie der vorliegenden Arbeit wurden die Studienteilnehmer in zahlreichen vorherigen Studien in ihrer Strategiewahl eingegrenzt (e.g., Lemaire & Brun, 2014, 2016; Lemaire & Lecacheur, 2011). Während Kinder zwischen der Abrundungsstrategie und der Aufrundungsstrategie wählen durften, wurde die Anwendung der gemischten Rundungsstrategie (d.h. einen Operanden ab- und den anderen aufzurunden) ausgeschlossen. Obwohl dies ein bewusst gewähltes und legitimes Studiendesign darstellt, liegt die Vermutung nahe, dass es die ökologische Validität der Ergebnisse einschränkt. Da Teilnehmer die gemischte Rundungsstrategie in einem natürlichen Setting ebenfalls anwenden (Xu et al., 2014), könnte die Einschränkung der Strategiewahl abweichende kognitive Anforderungen bei der Strategieanwendung erfordern. Aus diesem Grund wurde in Studie 2 der Strategiegebrauch von Kindern untersucht, ohne die gemischte Rundungsstrategie auszuschließen. Ferner sollte überprüft werden, welche Rolle exekutive Funktionen beim Strategiegebrauch einnehmen, wenn die Nutzung der gemischten Rundungsstrategie erlaubt ist. Hierzu bearbeiteten 88 Viertklässler Überschlagsaufgaben. Überdies wurden Aufgaben zu den exekutiven Funktionen Arbeitsgedächtnis sowie Shifting in die Analysen aufgenommen. Die Ergebnisse zeigten, dass Kinder zuverlässig Summanden mit Einerzahlen kleiner als fünf abrundeten und Summanden mit Einerzahlen größer als fünf aufrundeten, um Überschlagsergebnisse zu erhalten. Somit lieferten die Ergebnisse keinen Anhalt darauf, dass Kinder beide Einerzahlen der Summanden simultan in die Strategiewahl einfließen ließen oder Kompensationsstrategien anwendeten, um Abweichungen im Rundungsprozess auszugleichen (bspw. die gemischte Rundungsstrategie bei einer Aufgabe wie 73 + 24 anzuwenden). Ferner konnte kein Einfluss der untersuchten exekutiven Funktionen auf den Strategiegebrauch der Kinder gefunden werden. Die Ergebnisse stellen somit die ökologische Validität von bisherigen Studien mit eingeschränkter Strategiewahl in Frage und deuten darauf hin, dass die Rolle von exekutiven Funktionen in bisherigen Studien durch das Strategiedesign möglicherweise überschätzt wurde.

Die Einschränkung des Strategiegebrauchs wurde von vielen Autoren darin begründet, die Aufgabenschwierigkeit zu erhöhen sowie Deckeneffekte zu verhindern (Lemaire & Brun, 2014, 2016; Lemaire & Lecacheur, 2011). Letztlich wurde jedoch bislang in keiner Studie untersucht, ob Kinder tatsächlich eine bessere Strategiewahl aufweisen, wenn sie zwischen drei statt nur zwei Strategien wählen dürfen. Auch bleibt offen, ob die konsistent gefundenen Leistungsunterschiede zwischen homogenen und heterogenen Problemen lediglich durch das eingeschränkte Strategieset artifiziell hervorgerufen wurden. Diesbezüglich deuten die Ergebnisse von Studie 2 bereits darauf hin, dass bei einem Design mit drei verfügbaren Strategien die gemischte Rundungsstrategie zuverlässig bei heterogenen Problemen angewendet wird. Daher wurde in Studie 3 die Strategiewahl von Kindern unmittelbar gegenübergestellt, wenn diese zwischen drei Strategien (Abrundungsstrategie, Aufrundungsstrategie und gemischte Rundungsstrategie) oder lediglich zwei Strategien (Abrundungsstrategie und Aufrundungsstrategie) wählen konnten. 449 Drittklässler und 276 Viertklässler wurden angeleitet, die beste Strategie für Überschlagsaufgaben zu wählen. Im ersten Durchgang standen ihnen drei Strategien zur Verfügung (Drei-Strategie Bedingung), im zweiten Durchgang zwei Strategien (Zwei-Strategie Bedingung). Ferner wurde die Updating Leistung der Kinder erhoben. Die Ergebnisse zeigten, dass Kinder mit höherer Updating Leistung die beste Strategie gleichermaßen oft für homogene wie heterogene Aufgaben in beiden Bedingungen wählten. Kinder mit niedriger Updating Leistung hingegen zeigten eine schlechtere Strategiewahl für heterogene Aufgaben in beiden Bedingungen. Dies lässt darauf schließen, dass Updating eine zentrale Rolle spielt, sobald Instruktion oder Aufgabenmerkmale verschiedenartige Rundungsprozeduren innerhalb einer Aufgabe erfordern.

In zahlreichen vorherigen Studien zum arithmetischen Strategiegebrauch wie auch den vorherigen Studien der vorliegenden Arbeit wurden die Überschlagsrechenaufgaben visuell und ohne Zeitbegrenzung dargeboten (LeFevre et al., 1993; Lemaire & Lecacheur, 2011). Im schulischen Kontext sowie im alltäglichen Leben hingegen variiert die Präsentationsform von Aufgaben oftmals. So können Aufgaben mündlich statt schriftlich oder kurzzeitig statt permanent vorgegeben werden. Studien konnten bereits zeigen, dass arithmetische Leistungen von der Präsentationsform beeinflusst werden (Fürst & Hitch, 2000; LeFevre, Lei, Smith-Chant, & Mullins, 2001). Unklar ist jedoch, inwiefern dies mit Variationen im Strategiegebrauch sowie kognitiven Voraussetzungen der Teilnehmer in Verbindung steht. Um dies zu untersuchen, wurden in Studie 4 Dritt- und Viertklässler zufällig zu einer von drei Präsentationsbedingungen zugeteilt: einer unbegrenzt visuellen, einer zeitbegrenzt visuellen sowie einer auditiven Präsentationsbedingung. In Experiment 1 (207 Drittklässler und 140 Viertklässler) bearbeiteten Kinder Überschlagsaufgaben und sollten zwischen der Abrundungs-, Aufrundungs- und gemischten Rundungsstrategie wählen. In Experiment 2 (277 Drittklässler und 136 Viertklässler) bearbeiteten Kinder die gleichen Überschlagsaufgaben, bei welchen nun die beste Strategie angezeigt wurde. Die Ergebnisse zeigten, dass Kinder die besten Leistungen hinsichtlich Strategiewahl und -ausführung bei unbegrenzt visueller Präsentation erzielten. Dies lässt sich am besten damit erklären, dass die visuelle Aufgabe als permanenter, externer Input dient, auf welchen die Kinder während des Rechenprozesses zurückgreifen können. Interessant ist, dass Kinder mit besseren Updating Leistungen sogar in erhöhtem Maße von dieser Präsentationsform profitierten. Dies spricht dafür, dass diese Kinder die Vorteile der externen Aufgabenpräsentation in effektiverer Weise nutzen konnten, indem sie beispielsweise die Aufgabe bei Bedarf erneut enkodierten oder Ergebnisse kontrollierten.

Zusammenfassend lässt sich festhalten, dass verschiedene kognitive Prozesse bei Überschlagsrechnungen beteiligt sind und den Strategiegebrauch von Kindern beeinflussen. Die Untersuchung individueller Unterschiede in exekutiven Funktionen sowie die Variation verschiedener Aufgabenparameter im Rahmen von vier Studien der vorliegenden Arbeit konnte differenziertere Einblicke in ebenjene Prozesse liefern. So konnte aufgezeigt werden, dass individuelle Unterschiede im Arbeitsgedächtnis Updating einen zentralen Einfluss auf den Strategiegebrauch zu nehmen scheinen. Interaktionen von Updating mit verschiedenen Aufgabenvariationen (Problemgröße, Problemtypen, Zahl der verfügbaren Strategien, Präsentationsform) deuten darauf hin, dass individuelle, kognitive Voraussetzungen der Kinder mit verschiedenartigen Prozessen in Bezug auf das Überschlagsrechnen zusammenhängen. In Zusammenschau der Befunde kann davon ausgegangen werden, dass Updating für die Konsolidierung sowie den Abruf von Strategien, die Gewichtung relativer, assoziierter Kosten von Strategieprozeduren, die Enkodierung von Aufgaben sowie den Rechenprozess eine zentrale Rolle spielt.

ABSTRACT

Computational estimation is an important skill in everyday life as well as in educational contexts. In the last decades, research has found that children use several strategies in computational estimation and that children's strategy use depends on different parameters. Still, little is known about the underlying cognitive processes. In the present work, we addressed this issue by investigating (1) the influence of individual differences in children's executive functions on their strategy use and (2) the influence of varying specific task and problem characteristics that are discussed to involve different cognitive processes.

In four studies, we asked third and fourth graders to solve computational estimation tasks by rounding the summands. Study 1 addressed the influence of working memory updating. The study found that efficient updating contributed to children's strategy use and moderated relations with problem characteristics. A deliberate feature of Study 1 was to restrict participants' strategy choice to the rounding-down and rounding-up strategies. Study 2 in turn investigated children's strategy use when mixed-rounding was allowed. Results indicated that children did not consider unit digits of both operands jointly. Also, no influence of executive functions could be found. Consequently, in Study 3, children's strategy selection when they could choose between three versus only two strategies was contrasted and the role of working memory updating was investigated. Indeed, children chose the best available strategy more often when three strategies were available. Importantly, relative strategy selection performance differed with children's updating capacities. Finally, Study 4 addressed another task variation that is important in everyday life and educational contexts. That is, presentation duration and modality were varied. Data showed that a permanent, written format was most beneficial for children's strategy use and that children's updating moderated presentation effects.

In sum, the results of the present work could shed some light onto cognitive processes in children's strategy use in computational estimation. Specifically working memory updating seems to contribute to third and fourth graders strategy use. Interpreting interactions with different task variations, updating most likely influences associative processes, long-term memory consolidation and retrieval as well as encoding and calculation processes.

INTRODUCTION

Strategies are pervasively used to help us cope with cognitive demands in everyday life. Take the example of a person using memory strategies to be able to remember the names of groceries he or she wants to buy in the supermarket without the need of a shopping list. As another example, it might be helpful for that person to use computational estimation to get an approximate sense of how much he or she must pay for these groceries at checkout. One could come up with many more examples, specifically ones in which estimation serves as a very helpful skill in everyday situations. Besides the importance in everyday life, computational estimation is a crucial component of children's mathematical cognition as it is not only related to other arithmetic skills but also to general measures of mathematical ability, such as arithmetic achievement scores (Dowker, 2012). Being able to correctly and efficiently perform computational estimations allows children to make approximate calculations without the need for a calculator or paper and pencil, for instance, to check the reasonableness of complex calculations found through other means. In addition, it provides information about children's general understanding of mathematical concepts and might help them to "develop a better understanding of place value, mathematical operations, and general number sense" (Star, Rittle-Johnson, Lynch, & Perova, 2009, p. 569).

Despite the omnipresence of estimation in children's lives and educational contexts, far less is known about cognitive processes being involved in estimation than about other basic numerical processes, such as counting and simple arithmetic (Dowker, 2012; Geary, 1994; Siegler & Booth, 2005). The present work seeks to provide some new insights into these cognitive processes of children's strategy use in computational estimation. For this purpose, in four studies, we[2] (1) investigated the influence of individual differences in executive functions on children's strategy use and (2) varied different task parameters (i.e., strategy set, presentation condition) and problem characteristics (i.e., problem type, problem size) which are discussed to influence different cognitive processes during estimation and interact with individual differences.

[2] Following the practice of empirical studies, I employ '*we*' instead of '*I*' for the entire work.

1 THEORETICAL BACKGROUND

1.1 Strategy Use in Computational Estimation

Computational estimation refers to “the process of simplifying an arithmetic problem using some set of rules or procedures to produce an approximate but satisfactory answer through mental calculation” (LeFevre, Greenham, & Waheed, 1993, p. 95). Computational estimation is an important component of mathematical cognition and provides information about the understanding of other mathematical concepts and about the interplay between different mathematical procedures (e.g., Bisanz & LeFevre, 2013; Lemaire, Arnaud, & Lecacheur, 2004; Siegler & Booth, 2005; Sowder & Wheeler, 1989; Star et al., 2009). According to Case and Sowder (1990), the estimation of addition and multiplication problems consists of two steps: (a) Converting the original, exact numbers of the problem to approximate numbers and (b) using these approximate numbers to mentally calculate the estimate. Clearly, there are numerous ways or strategies to convert the exact numbers of an arithmetic problem to approximate numbers. Depending on these strategies, estimations differ in their efficiency, accuracy, and complexity of used procedures.

A strategy can be defined as ‘‘a procedure or set of procedures for achieving a higher level goal or task” (Lemaire & Reder, 1999, p. 365). In the context of computational estimation, there have been different approaches to cluster these procedures into broader categories. Reys, Rybolt, Bestgen, and Wyatt (1982) identified three general families of people’s computational estimation strategies: reformulation, compensation, and translation. Reformulation refers to changing the numbers through rounding (e.g., rounding 47 to 50) or truncation of nonsignificant digits (e.g., truncating 47 to 40). Compensation refers to adjusting the numbers used to produce the estimate (prior compensation) or the produced estimate itself (post compensation) to reduce discrepancies between estimates and exact calculations. That is, “compensation might involve rounding one number up and one number down to produce an estimate that stays relatively close to the actual answer (e.g., 45 x 65 to 40 x 70)” (LeFevre et al., 1993, p 97). Hence, compensation might involve rounding procedures as in translation strategies but emphasizes the goal of proximity over simplicity of estimates. Translation refers to changing the mathematical structure to create a mentally more manageable problem (e.g., translating 83 + 74 + 82 to 3 x 80). The two latter categories are considered to be more sophisticated approaches and to reflect conceptual understanding of estimation procedures (Reys et al., 1982). Although this sophisticated use and full conceptual knowledge of estimation strategies seem not to emerge till early adolescence (Case & Sowder, 1990; LeFevre et al., 1993; Sowder & Wheeler, 1989), some estimation capacities emerge very early (Dowker, 1997, 2012). Indeed, children as young as ten years-old regularly use rounding strategies to estimate sums or products of arithmetic problems (LeFevre et al., 1993; Lemaire, Lecacheur, & Farioli, 2000).

As children clearly preferred rounding the operands to simplify arithmetic tasks over other strategies (LeFevre et al., 1993; Lemaire et al., 2000), studies in children have mainly focused on the use of rounding strategies in computational estimation. Most studies agree that children typically use multiple strategies and that they adapt their strategy choice to features of the estimation problem (e.g., LeFevre et al., 1993; Lemaire & Lecacheur, 2002). At the same time, individuals differ in how well they adapt their strategy use to problem features and how fast they solve estimation problems. With increasing age, children tend to rely less on easier strategies, select the best strategy for a problem more often, and execute strategies more efficiently (for overviews, see Lemaire, 2018; Siegler, 2007). Moreover, there is evidence that different problem features, such as problem size (i.e., size of the exact calculation) and problem type (i.e., different combinations of unit digits), influence participants' strategies and performance in arithmetic (LeFevre et al., 1993; Lemaire et al., 2000; Lemaire & Lecacheur, 2011). Most of these findings are explained in terms of certain cognitive processes being involved in computational estimation. The present work focuses on these cognitive processes and seeks to give some insights into the role of individual differences and variations of task and problem characteristics in children's computational estimation strategies.

1.2 Cognitive Processes in Arithmetic Strategy Use

Different cognitive processes are involved when participants use strategies to accomplish arithmetic tasks. For example, estimating sums of addition problems first requires encoding the original problem. As a next step, associative processes influence which strategy is selected. That is, theoretical models of strategy use (Lovett & Anderson, 1996; Rieskamp & Otto, 2006; Shrager & Siegler, 1998) propose that relative benefits and costs of available strategies are activated and that the strategy with larger benefits and lower costs is selected. According to the selected strategy, operands are then manipulated (e.g., by rounding the summands), manipulated numbers are held in working memory, and simple calculation procedures occur to add the manipulated operands. In addition, participants must disengage from the just executed strategy and above-mentioned strategy selection procedures are reactivated to estimate the sum of a following trial. How well these lower-order cognitive processes are accomplished strongly depends on task, problem, and participant characteristics.

Research showed that proficiency in arithmetic strategy use is positively correlated with intelligence (Reys et al., 1982) and other components of mathematical cognition, such as arithmetic fluency and simple calculation performance (Dowker, 1997, 2012; LeFevre et al., 1993; Seethaler & Fuchs, 2006). Only recently, interest in other cognitive correlates, such as working memory or executive functions, has grown rapidly. Individual differences in these cognitive correlates are discussed to influence participants' capacities while accomplishing strategy procedures and to operate the above-mentioned lower-order cognitive processes of strategy use. Another perspective on investigating cognitive processes in strategy use is to vary task or problem characteristics that are discussed to place cognitive demands on executive functions and to involve different cognitive processes. We briefly review the literature on these two perspectives before outlining the logic of the four studies within the present work.

1.2.1 Individual Differences in Executive Functions

In addition to basic arithmetic capabilities (Dowker, 1997, 2012; LeFevre et al., 1993; Seethaler & Fuchs, 2006), executive functions (EF) are discussed to have major impact on arithmetic performance in general and arithmetic strategy use in particular. EF are higher-order cognitive processes that operate other cognitive processes during goal-directed behavior (Friedman & Miyake, 2017). With its historical roots in neuropsychological studies, EF are often related with functions of the frontal lobes (Miyake et al., 2000). For many years, executive control processes have been associated with the central executive in the working memory model proposed by Baddeley and Hitch (1974). However, later works suggested that rather than being a unitary system, the central executive can be divided further into different subprocesses (Baddeley, 1996; Karbach & Kray, 2016; Miyake et al., 2000). These subprocesses include inhibition, shifting, and working memory updating (for overviews, see Diamond, 2013; Miyake et al., 2000). Inhibition refers to suppressing a dominant response and resisting distractions. Shifting is the ability to switch between different tasks or cognitive sets.

Working memory updating refers to monitoring and revising information held in working memory. Analyses regarding the structure of these three core EF showed that they correlate with each other and share some common variance but at the same time resemble differentiable cognitive processes (Duncan, 2010; Fisk & Sharp, 2004; Miyake et al., 2000) and can also be differentiated from general measures of working memory capacity. As an example, updating skills and working memory capacity are closely linked to one another as both involve temporarily storing items in working memory. In contrast to measures of working memory capacity, updating additionally involves manipulating and substituting information held in working memory which may independently contribute to predicting higher mental abilities (Ecker, Lewandowsky, Oberauer, & Chee, 2010).

Numerous studies have been conducted on the influence of EF on mathematical abilities (for overviews, see Bull & Lee, 2014; Cragg & Gilmore, 2014). Studies consistently found working memory updating to be a unique predictor of arithmetic achievements. Findings regarding the influence of inhibition and shifting on the other hand are less conclusive. Lechuga, Pelegrina, Pelaez, Martin-Puga, and Justicia (2016) investigated the relative contributions of working memory updating and intelligence to academic attainment in fourth graders. In hierarchical regression analyses, they found that updating accounted for a larger amount of variance in the prediction of mathematical problem solving and arithmetical operations than children’s intelligence. Similarly, Yeniad, Malda, Mesman, van IJzendoorn, and Pieper (2013) investigated the association of shifting and intelligence with math performance in a meta-analysis. They found that both predictors were related to math performance. However, intelligence demonstrated stronger associations with math performance than shifting abilities. Regarding unique contributions of single EF on arithmetic achievements, a meta-analysis by Friso-van den Bos, van der Ven, Kroesbergen, and van Luit (2013) revealed a higher correlation between mathematics and updating than between mathematics and inhibition or shifting. Similarly, St Clair-Thompson and Gathercole (2006) identified high unique associations of a factor including updating and working memory span measures with mathematics scores in a principal component analysis, whereas inhibition only accounted for a small amount of unique variance and no shifting factor could be identified. Another confirmatory factor analysis in second graders even found that inhibition and shifting did not predict arithmetic at all, when controlling for updating (van der Ven, Kroesbergen, Boom, & Leseman, 2012).

The previous findings imply that EF are crucially involved in arithmetic performance. However, the role of EF in arithmetic strategy use is remarkably less well investigated. Only one study by Lemaire & Lecacheur (2011) investigated the influence of children's individual differences in inhibition and shifting on their computational estimation strategy use. They found that children with better inhibition and shifting chose the best strategy more often and executed strategies more efficiently. In addition, other findings imply that working memory processes are involved in arithmetic strategy use. That is, Barrouillet and Lépine (2005) found that children with high working memory spans tended to use more sophisticated strategies more often than children with low working memory spans. Also, loading the central executive with a secondary task led children to obtain worse strategy selection performance (Ai, Yang, Zhang, Si, & Liu, 2017; Imbo & Vandierendonck, 2007b). Hence, there is first evidence that EF are involved in arithmetic strategy use. It is discussed that this occurs via influencing above-mentioned lower-order processes of strategy use. That is, EF should be involved in strategy selection processes by activating available strategies in working memory and choosing among strategies as a function of problem characteristics by retrieving associations between strategies and certain problem features from long-term memory (Ai et al., 2017; Barrouillet & Lépine, 2005; Lemaire & Lecacheur, 2011). In addition, EF should be relevant in strategy execution when storing and retrieving intermediate results in and from working memory is required (Fürst & Hitch, 2000; Imbo & Vandierendonck, 2007b).

1.2.2 Variations in Task and Problem Characteristics

As mentioned before, several task and problem characteristics influence participant's strategy selection and execution procedures. Therefore, another perspective on investigating cognitive processes involved in strategy use is to experimentally vary task and problem features and interpret performance differences under these variations.

As an example, previous research found that more complex problems place larger cognitive demands on working memory and negatively influence mental calculation processes as a consequence. That is, addition problems with a carry over the hundreds are associated with less accurate calculations (DeStefano & LeFevre, 2004; Kalaman & LeFevre, 2007) and the addition of larger numbers is associated with slower calculation times (for an overview on the so-called problem size effect, see Ashcraft, 1992). Consequently, it is assumed that the size of a problem influences mental calculation processes within arithmetic tasks.

Also, another problem characteristic that has been investigated in participants' computational estimation strategy use is the size of unit digits. According to the unit digits of the operands, arithmetic problems can be divided into different problem types. That is, problems with operands' unit digits both smaller or both larger than five are called homogeneous problems (e.g., 32 + 21 or 28 + 39) and problems with one operand's unit digit smaller and the other operand's unit digit larger than five are called heterogeneous problems (e.g., 32 + 29). Many studies found that participants' strategy performance was better for homogeneous than heterogeneous problems (Hodzik & Lemaire, 2011; Lemaire & Brun, 2016; Lemaire & Lecacheur, 2011). Importantly, these problem type effects were consistently found in studies with a restricted strategy set. That is, participants were told to choose between the rounding-down strategy (i.e., rounding both operands down to the closest decades) and the rounding-up strategy (i.e., rounding both operands up to the closest decades). This restriction was in place even for heterogeneous problems for which mixed-rounding (i.e., rounding one operand down and the other up to the closest decades) would yield the closest estimate. Consequently, strategy selection differences for the different problem types can be best explained in terms of above-mentioned associative processes involved in selection procedures. That is, associations between strategies and certain problem features are strengthened as a result of learning experience and repeated use of algorithms (Siegler & Shrager, 1984). Therefore, only for homogeneous problems children could form stable associations between problem features (i.e., unit digits) and the available strategies (i.e., rounding-down and rounding-up strategies), whereas heterogeneous problems would be associated with the mixed-rounding strategy, a strategy that was not allowed in studies with a restricted strategy set. Consequently, this implies that the set of available strategies within a task is another crucial component of cognitive processes associated with strategy selection.

Additionally, different rounding strategies involve procedures of varying complexity. Regarding cognitive costs, the rounding-up strategy requires more complex processes (i.e., rounding up the first operand, holding the rounded decade in working memory, executing the same procedure for the second operand, and adding the two updated decades) than the rounding-down strategy (i.e., solely adding the decade units). Therefore, the rounding-up strategy is associated with relatively larger costs in associative processes of the selection phase than the rounding-down strategy (Lemaire & Lecacheur, 2002).

Finally, variations in the presentation of arithmetic problems have been found to influence processes within participants' strategy use. As an example, Lemaire and Brun (2014) showed that children selected the better strategy more often under presentation with longer response stimulus intervals (i.e., duration between participant's response and next problem display) than shorter response stimulus intervals. With longer intervals children were better able to disengage from the just-executed strategy, reactivate the available set of strategies and switch strategies when appropriate. As another example, adults selected the better strategy more often under a long than under a short presentation duration (Lemaire & Brun, 2016). This was explained in terms of longer presentation facilitating encoding processes to construct a more precise representation of the problem in working memory. Moreover, not only presentation duration but also presentation format and modality have been discussed to influence the construction of mental representations of the task. That is, many models of information processing (e.g., Baddeley, 1992; Campbell, 1992, 1994; Mayer, 2014; Paivio & Lacey, 2007) share the assumption of two separate channels for verbal and nonverbal material or auditory and visual input. Consequently, it is assumed that both involve different representational codes: the activation of phonological codes for aurally presented problems and visual codes for visually presented problems. Prior works found that participants made more errors in simple arithmetic when problems were presented aurally or in number word format rather than as visual digits (Adams & Hitch, 1997; Campbell, 1994; Klingner, Tversky, & Hanrahan, 2011; LeFevre, Lei, Smith-Chant, & Mullins, 2001). The authors concluded that phonological codes activated by auditory format interacted with calculation or answer production phases of arithmetic. However, unknown is whether variations in presentation modality would also be accompanied by systematic variations in children's strategy use, indicating that representational codes would also interact with strategy selection phases, such as the activation of associated costs and benefits.

In sum, various problem and task characteristics seem to influence cognitive processes that are involved in children's strategy use when solving computational estimation tasks. According to the literature, presentation characteristics most likely influence encoding processes of the task. In turn, available strategy sets and problem types as well as the complexity of strategy procedures most likely affect associative processes in the strategy selection phase. Finally, the problem size seems to influence calculation processes. The main goal of the present work is to investigate the influence of varying such problem and task characteristics on children's strategy use in computational estimation and how postulated effects of problem and task features would interact with individual differences in children's EF.

2 THE PRESENT WORK

The present work seeks to shed some light onto cognitive processes in children's strategy use in computational estimation. For that purpose, we investigated the contributions of individual differences in EF to children's strategy use and experimentally varied task and problem characteristics that are discussed to involve different cognitive processes. In the following, we briefly outline the theoretical basis of the empirical studies and the research questions therein.

In the past decades, numerous studies have been conducted on parameters that would influence participants' strategy use in computational estimation. Experimental effects, such as effects of dual-task paradigms (Duverne, Lemaire, & Vandierendonck, 2008; Imbo & Vandierendonck, 2007a; Logie, Gilhooly, & Wynn, 1994) and sequential effects (Lemaire & Brun, 2014; Uittenhove & Lemaire, 2013), have been explained in terms of cognitive demands being placed on working memory during computational estimation procedures. Also, direct effects of inhibition and shifting (Lemaire & Lecacheur, 2011) on children's strategy use in computational estimation imply that EF contribute to strategy selection and execution. Despite this evidence, the role of one EF, namely working memory updating, has not yet been investigated in children's strategy use. Therefore, **Study 1** addressed this issue and investigated the contributions of updating to children's strategy selection and execution. We hypothesized that computational estimation procedures such as storing and retrieving intermediate results in and from working memory should be facilitated in children with more efficient updating processes. In addition, updating should be involved when activating available strategies in working memory and choosing among strategies as a function of problem characteristics by retrieving associations from long-term memory. Hence, the research questions of **Study 1** were:

- Does working memory updating influence children's strategy selection and execution?
- Does working memory updating moderate effects of problem type (i.e., homogeneous/ heterogeneous problems) and problem size (smaller/ larger problems)?

In Study 1 as in a large number of previous studies on strategy use in computational estimation, participants were constrained in their strategy use by only allowing the use of the rounding-down and the rounding-up strategy but excluding the use of mixed-rounding (e.g., Lemaire & Brun, 2014, 2016; Lemaire & Lecacheur, 2010, 2011). While this is a deliberate and legitimate experimental study design, it most likely limits the ecological validity. A few studies in which participants were free in their strategy choice (e.g., LeFevre et al., 1993; Xu, Wells, LeFevre, & Imbo, 2014) found that participants use the mixed-rounding strategy roughly as often as rounding-down and rounding-up. However, there has been no detailed investigation of the problem types children would use the mixed-rounding strategies on. That is, unknown is whether children would use the mixed-rounding strategy adaptively according to the operands' unit digits or whether the use of the mixed-rounding strategy would indicate that children consider both unit digits jointly and opt for compensation. In **Study 2** we inspected children's strategy use on different problem types (i.e., small-unit, large-unit, and mixed-unit problems) to gain a more detailed picture of children's cognitive selections mechanisms. Some of the previous studies with a restricted strategy set could reveal effects of EF on participants' strategy use (e.g., Lemaire & Lecacheur, 2011; Uittenhove & Lemaire, 2013). However, the restricted strategy set might have involved larger cognitive demands and exaggerated the role of EF. Therefore, **Study 2** additionally examined the role of EF on children's strategy use when mixed-rounding is allowed. The research questions in **Study 2** were:

- Do children use the mixed-rounding strategy according to the optimal problem-based strategy (i.e., on mixed-unit problems) or do they opt for compensation by considering both unit digits jointly?
- Do EF contribute to children's strategy selection and execution when mixed-rounding is allowed?

As outlined before, in numerous previous studies on strategy use in computational estimation, participants' strategy selection was restricted to the rounding-down and the rounding-up strategies (e.g., Lemaire & Brun, 2014, 2016; Lemaire & Lecacheur, 2010, 2011). The use of mixed-rounding was excluded to increase task difficulty and avoid ceiling effects. Better strategy performance on homogeneous problems than on heterogeneous problems was consistently found with this study design. However, studies are missing on whether indeed children are better in selecting the best available strategy when they can choose between three strategies than when they can choose between two strategies only. Also, no study has yet actually investigated whether problem type effects would still occur when three strategies are available. In **Study 3**, we addressed this existing gap in strategy research and documented the contribution of the number of strategies (two versus three strategies) to children's strategy selection. Also, as it is argued that restricting strategy selection to two strategies places larger cognitive demands on participants cognitive resources than when allowing to use the mixed-rounding strategy we investigated the moderating influence of working memory updating on task and problem type effects. In sum, the research questions of **Study 3** were:

- Do children obtain better strategy selection performance under a three-strategy condition than under a two-strategy condition?
- Do problem type effects (i.e., better performance on homogeneous than heterogeneous problems) occur both under the three-strategy condition and the two-strategy condition?
- Does working memory updating moderate effects of task condition and problem type?

Besides restricting children in their strategy use, another common study feature has been to present participants with problems visually and with no time restriction (e.g., LeFevre et al., 1993; Lemaire & Lecacheur, 2011). However, children are often confronted with cognitive tasks of various presentation formats, modalities, and durations in everyday situations and educational contexts. That is, problem display duration might be limited, or presentation of problems might be aural rather than visual. A few previous studies suggest that arithmetic performance varies under different presentation modalities and durations (e.g., Fürst & Hitch, 2000; LeFevre et al., 2001). Still, unknown is whether varying presentation conditions lead children to use different strategies and to obtain different levels of strategy performance. In addition, presentation effects have been theoretically explained in terms of placing different cognitive demands on participants' working memory or influencing participants' efficiency of encoding procedures. However, unknown is whether presentation effects are in fact moderated by individuals' characteristics, such as their efficiency of working memory updating. **Study 4** addressed these issues and directed the following research questions:

- Do presentation duration (i.e., time-unlimited/ time-limited) and modality (i.e., visual/ auditory) influence children's strategy selection and execution?
- Does working memory updating moderate effects of presentation conditions?

3 SUMMARY OF THE EMPIRICAL STUDIES

3.1 Study 1: Effects of Working Memory Updating on Strategy Use

Hammerstein, S., Poloczek, S., Lösche, P., Lemaire, P., & Büttner, G. (2019). Effects of working memory updating on children's arithmetic performance and strategy use: A study in computational estimation. *Journal of Experimental Child Psychology, 184,* 174-191.

Background. Decades of cognitive research found that EF are crucially involved in academic performance. In the field of arithmetic, specifically one EF, namely working memory updating, seems to be most strongly linked to task performance above and beyond the influence of other EF and even general intelligence (Friso-van den Bos et al., 2013; Lechuga et al., 2016; St Clair-Thompson & Gathercole, 2006; van der Ven et al., 2012). Despite this strong evidence on the influence of updating on arithmetic performance, no study has yet investigated the contributions of updating to participants' strategy use in computational estimation. Hence, the goal of the present study was to address this existing gap in strategy research. We asked third and fourth graders to estimate sums of two-digit addition problems (e.g., 42 + 76) with the rounding-down or the rounding-up strategy. We tested the hypothesis that children with more efficient updating would choose the better strategy more often and execute strategies more efficiently. Another goal was to examine whether the efficiency of updating moderated effects of problem characteristics, such as the size of problems and problem types.

Method. A total of 308 children were tested: The sample consisted of 158 third graders (90 males; age in months: M = 114.0, SD = 5.6, range = 102.4–138.7) and 150 fourth graders

(80 males; age in months: *M* = 126.1, *SD* = 5.2, range = 116.2–147.0). All participants completed a computational estimation task and four working memory updating tasks (spatial keep track, day keep track, frog position updating, and color updating tasks). In the computational estimation task, children were asked to estimate sums for two-digit addition problems. They were instructed to use one of two rounding strategies: either the rounding-down strategy or the rounding-up strategy. To investigate problem type effects, we included two types of problems: so-called homogeneous and heterogeneous problems.

Results. Regarding children's strategy flexibility, we identified two qualitatively distinct subgroups. The first subgroup consisted of children approaching the problems of a test block in a flexible manner and adjusting strategies on a problem-by-problem basis; the other subgroup showed an inflexible approach and solved all or nearly all problems with the same strategy, either rounding-down or rounding-up. Data revealed that more children with more efficient updating tended to use the two available strategies in a flexible manner. Because children's flexibility is necessarily linked to the accuracy of their strategy use, we focused on flexible test blocks in the analyses of children's better strategy selection and estimation latencies. Data showed that children with more efficient updating were more likely to select the better strategy on a problem-by-problem basis than children with less efficient updating. This relation was qualified by these children (a) using the rounding-up strategy more often and (b) being specifically adaptive on homogeneous problems. Moreover, we found that children with more efficient updating were specifically faster on (a) larger addition problems and (b) homogeneous problems.

Discussion. The current study found that efficient updating contributed to children's strategy selection and strategy execution and that relations between children's strategy use and updating changed with problem characteristics. Children with more efficient updating were more likely to approach the task in a flexible manner. This might be best explained by assuming that working memory updating is involved in the execution of processes crucial for trial-by-trial strategy selection (i.e., reactivating both strategies in working memory, analyzing problem features to determine which strategy is the best, and choosing the better strategy before executing it). Furthermore, more efficient updating led children to select the better strategy more often and execute strategies more efficiently specifically on larger problems, on familiar problem types, and by using the more complex strategy more often. These findings can be best explained in terms of the associative processes that are activated when being presented with a problem. For children with more efficient updating, more complex procedures and harder problems entail costs that are smaller in relation to their available updating resources. In addition, given the link between working memory and long-term memory (Barrouillet & Lépine, 2005; Unsworth, 2016, 2019), children with more efficient working memory updating likely have had an advantage in forming associations and in retrieving those established associations for familiar homogeneous problems. The present findings provide first evidence that updating contributes to children's strategy use and indicate that updating is involved in associative processes of strategy selection.

3.2 Study 2: Mixed-rounding Strategy Use

Poloczek, S., **Hammerstein, S.**, & Büttner, G. (2020). Children's mixed-rounding strategy use in computational estimation. *Manuscript submitted for publication*.

Background. A methodological feature of many recent studies on computational estimation (e.g., Ai et al., 2017; Lemaire & Brun, 2016; Hodzik & Lemaire, 2011; Lemaire & Lecacheur, 2010, 2011; Lemaire, Luwel, & Brun, 2017) was that participants' choice of strategies was restricted to rounding either both operands down (i.e., rounding-down strategy) or rounding both operands up (i.e., rounding-up strategy) to the nearest decades, even for problems in which mixed-rounding (i.e., rounding one operand up and the other down to the nearest decades) would yield the closest estimate. This restriction was a deliberate feature to increase difficulty of strategy selection and to avoid ceiling effects and no-variance in participants' strategy use. However, the restriction potentially limits the ecological validity of the obtained results: If participants preferred to choose mixed-rounding for mixed-unit problems, they'd have to inhibit this familiar choice and consider both operands' unit digits simultaneously to select the best out of the other available strategies. A few studies with no strategy restrictions found that participants used the mixed-rounding strategy roughly as often as rounding-down and rounding-up (e.g., LeFevre et al., 1993; Xu et al., 2014). However, unknown is whether children use the mixed-rounding strategy adaptively on different problem types with varying combinations of unit digits. Therefore, the present study investigated children's strategy choices on different problem types when mixed-rounding was allowed. We asked fourth graders to estimate sums of two-digit addition problems. We expected children to consistently use the rounding-down strategy on small-unit problems (i.e., unit digits both smaller than five), the mixed-rounding strategy on mixed-unit problems (i.e., one unit digit smaller and the other larger than five), and the rounding-up strategy on large-unit problems (i.e., unit digits both larger than five) with no systematic strategy selection differences within these main problem categories. In addition, we investigated how individual differences in best strategy use and estimation latencies varied with children's EF to assess whether previously revealed effects in studies with a restricted strategy set (e.g., Lemaire & Lecacheur, 2011; Uittenhove & Lemaire, 2013) might stem from artificially increased cognitive demands due to the exclusion of mixed-rounding.

Method. Eighty-eight fourth graders (46 males; age in months: M = 122.0, SD = 5.4, range = 111-138) participated in our study. They performed a computational estimation task, ten EF tasks, and arithmetic subskill tasks. In the computational estimation task, children were asked to estimate sums for two-digit addition problems. They were instructed to use one of three rounding strategies: either the rounding-down strategy, the rounding-up strategy, or the mixed-rounding strategy. Children's EF were assessed with three working memory tasks (spatial updating, color updating, and picture span backward tasks), three shifting tasks (animal color, color shape tasks and a trail making test), and four inhibition tasks (real animal size, object inhibition, flanker, and color stroop tasks). One shifting task and all inhibition tasks were excluded from further analyses due to a lack of reliability.

Results. Children varied in their strategy flexibility. Most children approached the problems of a test block in a flexible manner and adjusted strategies on a problem-by-problem basis; some children showed an inflexible approach and solved all or nearly all problems with the same strategy. Because children's flexibility is necessarily linked to the accuracy of their strategy use, we focused on flexible test blocks in the following analyses. As expected, flexible children's strategy selection was substantially influenced by main problem categories and did not depend on the exact sum of the unit digits within these main categories. That is, children almost always used rounding-down for small-unit problems, mixed-rounding for mixed-unit problems and rounding-up for large-unit problems. Children were most likely to select the best strategy on small-unit problems and were fastest when using the rounding-down strategy.

Individual differences in best strategy selection and estimation latencies could not be explained by differences in working memory or shifting and no switching costs were found.

Discussion. The present data revealed that children's strategy choices did not depend substantially on the exact sum of the unit digits but on the main problem categories. Hence, we found no evidence that children considered both unit digits jointly and applied prior or post compensation to reduce the rounding distortion of their rounding approach. This questions the ecological validity of studies examining computational estimation and excluding mixed-rounding by design as it indicates that the cognitive demands are different and higher than in a design without this restriction specifically on heterogeneous (i.e., mixed-unit) problems. Results regarding children's strategy selection and performance found that children were most accurate and fastest with the rounding-down strategy. This is best explained by mixed-rounding or rounding-up being more complex procedures as both require the mental manipulation of the problem during the approximation. No influence of EF, namely working memory and shifting, on children's strategy use and no switching costs could be found. The absence of these effects might indicate that the role of EF has been exaggerated in previous studies with a restricted strategy design (e.g., Lemaire & Lecacheur, 2011). Results of the present study imply that further research should be done with more realistic designs.

3.3 Study 3: Two Versus Three Strategies in Strategy Selection

Hammerstein, S., Poloczek, S., Lösche, P., & Büttner, G. (2020). Two versus three available strategies in children's strategy selection in a computational estimation task. *Manuscript submitted for publication.*

Background. In a large number of previous studies on strategy use in computational estimation (e.g., Ai et al., 2017; Lemaire & Brun, 2016; Hodzik & Lemaire, 2011; Lemaire & Lecacheur, 2011; Lemaire et al., 2017), participants were constrained in their strategy use by only allowing the use of the rounding-down or the rounding-up strategy, while the use of the mixed-rounding strategy was excluded to increase difficulty of strategy selection and to avoid ceiling effects. Studies with this strategy design consistently found that participants selected the best available strategy more often on so-called homogeneous problems than on heterogeneous problems (Hodzik & Lemaire, 2011; Lemaire & Brun, 2016; Lemaire &

Lecacheur, 2011). This was explained by higher cognitive demands for strategy selection on heterogeneous problems. However, no study has yet actually compared participants' selection performance when allowing to choose between three strategies to selection performance when allowing to choose between two strategies. Similarly, so far, no study investigated whether problem type differences would also occur under a three-strategy condition or disappear in the presence of mixed-rounding being allowed. Therefore, in the present study, we asked third and fourth graders to select the best strategy to estimate sums of two-digit addition problems (e.g., 42 + 76) in conditions, for which only two versus three strategies were available. We tested the hypothesis that children would choose the best strategy more often under the three-strategy than under the two-strategy condition. Another goal was to examine whether the efficiency of updating moderated effects of task condition and problem type. We expected children with more efficient updating to be less influenced by more difficult task and problem characteristics.

Method. A total of 725 children were tested: The sample consisted of 449 third graders (231 males; age in months: M = 112.7, SD = 6.8; range = 90-178) and 276 fourth graders (149 males; age in months: M = 124.5, SD = 5.9; range = 113-153). All children completed a strategy selection task under two task conditions and four working memory updating tasks (spatial keep track, day keep track, frog position updating, and color updating tasks). In the strategy selection tasks, children were asked to indicate the best available rounding strategy without actually calculating the estimates. First, children worked on problems under a *three-strategy* condition, in which they could choose between the rounding-down strategy, the mixed-rounding strategy, and the rounding-up strategy. On a second day, children worked on another set of problems under a *two-strategy* condition, in which they could choose only between the rounding-down and the rounding-up strategy.

Results. Data revealed that children were more likely to select the best available strategy when they could choose between three strategies than when they could choose between two strategies and children were more likely to select the best available strategy on homogeneous than heterogeneous problems. Children with more efficient updating were more likely to select the best available strategy than children with less efficient updating. Importantly, task condition and problem type effects were moderated by children's updating capacities. Children with less efficient updating clearly obtained worse performance on heterogeneous than on homogeneous problems both under two-strategy and three-strategy conditions. In contrast, children with more efficient updating obtained comparable selection performance for heterogeneous and homogeneous problems in both conditions.

Discussion. As expected, the present study found that children chose the best available strategy more often when three strategies were available than when only two strategies were available. Importantly, relative strategy selection performance differed with children's updating capacities. This indicates that children with more efficient updating could cope with the more demanding cognitive processes involved in selecting the best available strategy on heterogeneous problems when only the rounding-down and rounding-up strategies were available. In contrast, children with less efficient updating performed worse on heterogeneous problems not only under the two-strategy condition but also under the three-strategy condition. An explanation might be that these children could not yet create a stable association between heterogeneous problems and the mixed-rounding strategy. Working memory processes contribute to associations in long-term memory being established and to existing associations being accessed (Barrouillet, Bernardin, & Camos, 2004; Unsworth, 2016; Unsworth, Brewer, & Spillers, 2013). Therefore, children with less efficient working memory might have had difficulties in forming or retrieving associations for heterogeneous problems regardless of the number of available strategies. Another explanation might be that these children had difficulties in considering unit digits of both operands simultaneously and based their strategy selection mainly on the first or the second unit digit to save cognitive demands of holding both unit digits in memory during the selection process. The present findings stress that researchers in the field of strategies should be aware of the influence of strategy sets on participants' performance. In addition, the results indicate that it might be recommendable to adjust instructions or materials to account for the difficulties of children with less cognitive resources specifically on mixed-unit problems.

3.4 Study 4: Effects of Presentation Conditions on Strategy Use

Hammerstein, S., Poloczek, S., Lösche, P., Lemaire, P., & Büttner, G. (2020). Effects of presentation duration and modality on children's strategy selection and performance: A study in computational estimation. *Manuscript submitted for publication.*

Background. In the vast majority of studies on arithmetic, participants were presented with problems visually and with no time restriction (e.g., LeFevre et al., 1993; Lemaire & Lecacheur, 2011). While this is a legitimate study design in experimental research, children are often confronted with cognitive tasks of various presentation formats, modalities, and durations in real-life educational contexts. Some studies suggest that arithmetic performance varies under different presentation conditions (e.g., Fürst & Hitch, 2000; LeFevre et al., 2001). However, unknown is whether variations in presentation modality and duration are accompanied by systematic variations in children's strategy use and whether presentation effects are moderated by individual's cognitive capacities. Hence, the goal of the present study was to examine (1) the influence of presentation duration and presentation modality on children's strategy use and (2) the moderating influence of one EF, namely working memory updating, on presentation effects. In two experiments, we asked third and fourth graders to estimate sums of two-digit addition problems (e.g., 52 + 39). In Experiment 1, children could choose among three available strategies to investigate their *strategy selection*. In Experiment 2, the strategy to use was cued on each problem to investigate children's *strategy execution*. In both experiments, children were randomly assigned to one of three presentation conditions with a distinct combination of duration (i.e., time-limited/ time-unlimited) and modality (i.e., visual/ auditory). We expected children to show better strategy selection and execution when problems were presented with no time limitation than with time limitation and when problems were visually displayed than aurally presented. Furthermore, we expected presentation conditions to interact with individuals' updating capacities. Children with less efficient updating were expected to benefit more than children with more efficient updating from an unlimited, visual presentation.

Method. In Experiment 1, 347 children were tested: The sample consisted of 207 third graders (108 males; age in months: $M = 112.1$, $SD = 4.9$; range = 101-128) and 140 fourth graders (78 males; age in months: $M = 123.8$, $SD = 5.0$; range = 114-142). In Experiment 2, 363 children were tested: The sample consisted of 277 third graders (121 males; age in months: $M = 112.5$, $SD = 5.7$; range = 99-133) and 136 fourth graders (76 males; age in months: $M = 124.4$, $SD = 5.8$; range = 113-143). All participants completed the computational estimation task, an arithmetic fluency task, and four working memory updating tasks (spatial keep track, day keep track, frog position updating, and color updating tasks). In Experiment 1 (choice condition), children were asked to give an approximate answer for two-digit addition problems without actually calculating the exact sum. They were instructed to use one of three rounding strategies: the rounding-down strategy, the rounding-up strategy, or the mixed-rounding strategy. In Experiment 2 (no-choice condition), the strategy to use was cued for each problem. The cued strategy always matched the best strategy. To investigate presentation effects children were randomly assigned to one of three presentation conditions in both experiments: (1) a time-unlimited visual, (2) a time-limited visual, or (3) a time-limited auditory condition. In the time-unlimited visual condition, children saw each problem until they provided their answer. In the time-limited auditory condition, children were verbally presented with the addition problems via headphones. In the time-limited visual condition, children were visually presented with each problem for the same length as stimuli lasted in the auditory condition (i.e., 3000 ms).

Results. The sample consisted of two distinct subgroups of children's strategy flexibility with most children approaching the problems of a test block in a flexible manner and some children solving (almost) all problems with the same strategy. Data showed that children's strategy flexibility increased for children with more efficient updating as well as when problems were visually displayed with no time limitation and when presented aurally. Because children's flexibility is necessarily linked to the accuracy of their strategy use, we focused on flexible test blocks in the following analyses. Children were most likely to choose the best strategy and to execute strategies more accurately when problems were presented visually with no time limitation in contrast to when problems were presented aurally or displayed with time limitation. Children were fastest to estimate when problems were visually displayed with time limitation in contrast to time-unlimited or auditory presentation. Moreover, effects of presentation condition were moderated by updating. That is, best strategy selection increased most under time-unlimited presentation for children with more efficient updating.

Discussion. The two experiments showed that presentation conditions influenced children's strategy selection and strategy execution and that these presentation effects were moderated by individual differences in working memory updating. Children were fastest but also least flexible and executed strategies less accurately when problems were visually displayed with time limitation. The presence of time-pressure while being able to encode the task visually might have led children to engage in adaptive mechanisms to save time and cognitive resources. In contrast, children were most likely to select the best strategy and to execute strategies most accurately when problems were presented visually with no time limitation. Benefits from time-unlimited visual presentation are best explained by the problem display serving as a permanent, external input (Adams & Hitch, 1997), which participants can refer to throughout the trials, dispensing them from active maintenance of the problem representation within working memory. Interestingly, these effects were qualified by children's updating. That is, children with more efficient updating benefitted from time-unlimited visual presentation to select the best strategy even more than children with less efficient updating. This indicates that these children used the advantages of a continuous problem display in a more efficient manner. This could occur via several mechanisms. For example, children might more flexibly change the focus of their attention by re-encoding the task when needed, or by checking the plausibility of their estimates more often than children with less efficient updating. These findings imply that updating facilitated the efficiency of encoding processes and that varying task parameters might be used to improve children's academic performance in educational contexts.

4 GENERAL DISCUSSION

The aim of the present work was to gain new insights into cognitive processes involved in children's strategy use in computational estimation. For this purpose, in four studies, we (1) investigated the influence of individual differences in EF on children's strategy use and (2) varied different task parameters (i.e., presentation condition, strategy set) and problem characteristics (i.e., problem type, problem size) which are discussed to involve different cognitive processes during estimation and interact with individual differences.

4.1 Connections Between the Studies

The aim of Study 1 was to investigate the influence of one EF that has not been investigated before, namely working memory updating. Two types of problems (i.e., heterogeneous/ homogeneous) were included. Results showed significant contributions of updating to children's strategy use. Additionally, updating specifically facilitated strategy selection and execution on well-known homogeneous problems. In Study 1 as in numerous prior studies, children's strategy choice was restricted to rounding-down and rounding-up.

Consequently, Study 2 addressed children's strategy use when the mixed-rounding strategy was allowed. The goal was to investigate (a) whether children would use the three available strategies according to the problem's main categories (i.e., small-unit, mixed-unit, and large-unit), and (b) whether the role of EF might have been exaggerated in prior studies by restricting children's strategy choice. Findings showed that children in fact used the strategies consistently according to the problem's main categories and no influence of working memory and shifting on children's strategy use could be found. Comparing Studies 1 and 2, results imply that children's strategy use and the role of EF differ with the available strategy set. Also, Study 2 indicates that problem type differences between homogeneous and heterogeneous problems might disappear in the presence of mixed-rounding being allowed. However, as no study has actually compared children's strategy selection when three versus only two strategies are available, Study 3 contrasted these two conditions. Interestingly, not only did we find contributions of updating on strategy selection under both conditions but also interactions of task condition, problem type and updating. That is, children with more efficient updating showed comparable performance for heterogeneous and homogeneous problems, whereas children with less efficient updating showed worse performance on heterogeneous problems under both conditions. To account for another task variation that might affect the ecological validity of strategy tasks, Study 4 not only allowed the use of mixed-rounding but also varied presentation duration (i.e., time-limited/ time-unlimited) and modality (i.e., visual/ auditory) of the addition problems. Results showed that all children benefitted most from unlimited, visual presentation and children with more efficient updating did even more so.

There are some important connections between the studies that should be considered. Importantly, all four studies investigated the influence of individual differences in EF on children's strategy use in computational estimation. Studies 1, 3, and 4 focused on one specific EF, namely working memory updating. Study 2 included all three core EF (i.e., working memory, shifting, and inhibition[3]). In addition to investigating individual differences, various problem and task characteristics were varied: strategy sets, presentation conditions, problem types, and problem size. While in Study 1 only two strategies (i.e., rounding-down and rounding-up) were available, in Studies 2 and 4, the use of mixed-rounding was additionally

[3] Note that inhibition had to be excluded due to a lack of reliability.

allowed. Study 3 even contrasted the strategy sets by including task conditions either with three or two available strategies. Moreover, presentation modality and duration were varied systematically in Study 4. While in Studies 1, 2, and 3 problems were all presented visually with no time limitation, Study 4 contained additional conditions with time-limited visual or auditory presentation. Finally, problem characteristics were varied. Studies 1 and 2 both addressed the influence of problem size. In addition, Studies 1, 2, and 3 addressed the influence of problem types. Studies 1 and 3 both differentiated between homogeneous and heterogeneous problems while Study 2 made even more detailed distinctions between problem types according to the size of unit digits.

Regarding the influence of individual differences in children's EF on their strategy use, results of Studies 1, 3, and 4 consistently reported contributions of working memory updating on strategy selection and execution. In contrast, Study 2 could not find effects of working memory and inhibition on children's strategy use. Study 2 discussed that the absence of an effect of EF on strategy use might stem from an exaggerated influence of EF when strategies were restricted to rounding-down and rounding-up. However, this explanation could be refuted by the following studies. That is, in Studies 3 and 4, even under conditions with mixed-rounding being allowed effects of working memory updating were present. Also, Study 2 proposed to vary presentation conditions to increase problem difficulty and detect effects of EF more easily. Accordingly, Study 4 varied presentation modalities and durations. Still, updating contributed to children's strategy use under all presentation conditions and even more so in children selecting the best strategy when problems were presented visually with no time limitation as in Study 2. Consequently, we can only speculate about reasons for the absence of an effect of EF on children's strategy use in Study 2. Even though the studies did not differ in terms of problem difficulties on an item level (e.g., range of problem size), the relative problem difficulty might have been lower for children in Study 2, making it harder to detect effects of individual differences. Indeed, the remarkably faster estimation latencies of children in Study 2 (5 s) in comparison to fourth graders in Study 4 in the unlimited visual condition (12 s) might indicate that there is a substantial difference between the samples. Additionally, the tasks used to measure the working memory component of EF slightly varied between the studies and a factor for updating was extracted in Studies 3 and 4 but not in Study 2. That is, the component score of working memory in Study 2 might have been deteriorated by task-specific variance (e.g., differences in color processing, verbal, or spatial skills) which does not account for differences in children's strategy use. Also, Study 2 included a *picture span backward* task (i.e., reproducing sequences of drawings in reverse order) which requires the manipulation of stimuli only at the end of a trial. In contrast, the tasks used in the other studies focused on constantly updating information held in working memory throughout trials. Strategy use in a computational estimation task involves activating available strategies in working memory, removing all but the selected strategy procedures from memory, manipulating operands by rounding, and adding the manipulated operands. This series of processes likely require constant updating of information in working memory. Therefore, the purer measure of (constant) updating in Studies 1, 3, and 4 might have been more predictive for children's strategy selection and execution than the composite score in Study 2. Finally, effects of shifting and working memory might have been detected when a larger sample size would have been included in Study 2. In a synopsis of the present results, we can assume that updating plays a crucial role in children's strategy use. The role of shifting and inhibition on the other hand remains less clear. The present work shows that children were faster when they used the same strategy on (almost) all problems. This suggests that not switching between strategies is beneficial for children's strategy execution. At the same time, Study 2 could not find switching costs or contributions of shifting to children's strategy use in participants approaching the task in a flexible manner and choosing strategies on a trial-by-trial level. These findings imply that shifting skills are not crucially involved in strategy use when participants generally engage in strategy selection mechanisms on a trial-by-trial level. This might occur as the role of switching strategies over consecutive trials might be small in comparison to the multiple steps that must be carried out in a computational estimation task. Strategy selection mechanisms (i.e., activation of available strategies, choosing among strategies as a feature of problem characteristics et cetera) on the other hand seem to impose substantial burden on children's cognitive resources as seen in faster estimation latencies in inflexible children.

In addition to the influence of individual differences, effects of task and problem variations could be found. Both Studies 1 and 2 showed effects of problem size on children's strategy execution, indicating that the estimation of larger problems involves more time-consuming calculation processes. In addition, the interaction of working memory updating with problem size (Study 1) indicates that calculation processes (i.e., storing and retrieving intermediate results in and from working memory) are supported by children's efficiency of updating processes.

Furthermore, Studies 1, 2, and 3 investigated the influence of systematically varying unit digits of presented problems on children's strategy selection. Study 1 could find performance differences with better strategy selection on homogeneous than heterogeneous problems. Keep in mind that this is a common finding in studies with a restricted strategy set (Hodzik & Lemaire, 2011; Lemaire & Brun, 2016; Lemaire & Lecacheur, 2011). Consequently, Study 2 investigated the influence of problem types on strategy selection with the mixed-rounding strategy being allowed. The study not only differentiated between homogeneous and heterogeneous problems but documented a more detailed picture of children's strategy use when problems with different combinations of unit digits were presented. Results could find no evidence that children would consider unit digits of both operands jointly as would be required in strategy selection with a restricted strategy set as in Study 1. Additionally, findings imply that the difference between homogeneous and heterogeneous problems likely disappears with the mixed-rounding strategy being allowed. Study 3 then contrasted children's strategy selection with two versus three strategies being available. Interestingly, results differ somewhat from results of Studies 1. That is, with only two strategies available, children in Study 1 chose the better strategy less often for heterogeneous than homogeneous problems and for children with more efficient updating this difference was even larger. In contrast, in Study 3, children with more efficient updating selected the better strategy equally often for heterogeneous and homogenous problems. An important difference between the two studies is that Study 3 required strategy choice only, whereas Study 1 also required execution of the selected strategy. Additionally, the instructional focus in Study 3 was clearly on accuracy, whereas a balance of speed and accuracy was required in Study 1. Focusing the task-goal on accuracy of strategy use generally leads to better strategy selection performance (Xu et al., 2014). This indicates that children with more efficient updating can figure out the second-best strategy when only two strategies are available per se (Study 3) even though no association in long-term memory might be available. However, as this process is complex and time-consuming even children with more efficient updating might favor speed over accuracy more often when the instructional focus is also on speed of strategy execution.

Moreover, results of the studies indicate that efficient updating facilitates children to opt for strategies involving more complex procedures. Studies 1 and 4 showed that children with less efficient updating were more likely to approach the addition problems with an inflexible approach by using a single strategy on all or almost all problems. For most of the inflexible children the dominant strategy was the least complex rounding-down strategy. Similarly, Study 1 found that children with less efficient updating used the least complex rounding-down strategy more often even when they generally used the available strategies in a flexible manner. Additionally, Study 3 found that children with less updating showed difficulties when selecting the best strategy for problems with mixed unit digits. Hence, updating seems to facilitate the coordination of multiple stimuli in working memory when selecting strategies and the execution of strategies with more complex procedures.

To account for the fact that computational estimation is closely linked to exact computation (Dowker, 2012; LeFevre et al., 1993), Studies 2 and 4 assessed children's differences in arithmetic subskills in addition to individual differences in EF. Both studies showed that arithmetic fluency influenced children's strategy use. However, only Study 4 could find influence of EF when controlling for arithmetic skills. Possible explanations for the absence of this effect in Study 2 have been discussed previously.

Finally, it should be noted that all four studies addressed computational estimation strategies of children from Grade 3 and Grade 4. Studies 1, 3 and 4 included children from both age groups to investigate age-related differences, Study 2 investigated fourth graders only. We decided to investigate third and fourth graders following previous studies showing that important strategic changes occur in children of this age range (e.g., LeFevre et al., 1993; Lemaire & Lecacheur, 2011). Moreover, in German schools, where the data were collected, rounding rules and computational estimation are first introduced in third grade and further practiced in fourth grade. As expected, age-related improvements could be consistently found above effects of other individual differences in the present studies. This indicates that there are age-related changes other than the development of arithmetic skills and the investigated EF that account for some variance in children's computational estimation strategy use. These might be for example developmental changes in general intelligence, other arithmetic skills, or other components of EF.

In sum, results of the four studies emphasize the role of specifically one EF, namely working memory updating, in third and fourth graders' strategy use in computational estimation. Additionally, variations of task and problem characteristics indicate the involvement of different cognitive processes in computational estimation and that updating is involved in accomplishing these processes.

4.2 Implications for Practice and Future Research

The findings of the present work yield important implications on an empirical and theoretical basis. The studies were based on existing models of strategy selection (Payne, Johnson, & Bettman, 1993; Rieskamp & Otto, 2006; Shrager & Siegler, 1998) that share several core assumptions such as the activation of associated costs and benefits during strategy selection. However, these models do not make assumptions on the involvement of EF processes in a person's strategy use. Various works have already suggested that cognitive components, such as inhibition and shifting, should be added (Lemaire & Lecacheur, 2011; Luwel, Schillemans, Onghena, & Verschaffel, 2009). Results of the present work strengthen this case by showing that another EF, namely working memory updating, contributes to children's strategy selection and execution. Updating consists in ‘‘coding incoming information for relevance to the task at hand and then appropriately revising the items held in working memory by replacing old, no longer relevant information with newer, more relevant information" (Miyake et al., 2000, p. 57). Hence, updating is most likely involved in the activation of available strategies in working memory and the selection among strategies as a function of problem characteristics by retrieving associations from long-term memory. According to the findings on the contributions of updating on best strategy selection, updating could be included in existing models as influencing the relative importance or weights of associated costs and benefits to the person, similar to weighted additive heuristics (see also Payne et al., 1993). That is, associated costs due to the complexity of strategy operations might be compensated by efficient cognitive processes such as updating. This is also supported by the present findings on children's strategy flexibility. That is, children with more efficient updating processes were more likely to use available strategies flexibly rather than to opt for using a single strategy on (almost) all problems. Again, we can assume that updating processes facilitated the execution of processes crucial for trial-by-trial strategy selection (i.e., reactivating strategies in working memory, analyzing problem features to determine which strategy is the best, and choosing the best available strategy before executing it) alleviating the relative, cognitive costs associated with strategy selection processes. In addition, none of the existing models makes assumptions on strategy selection processes when different strategy repertoires are available in a task. Results of the present work indicate that restricting the number of strategies involves additional cognitive processes within strategy selection (e.g., inhibition of the dominantly associated strategy and selection of the best out of the available strategies by some alternative way) leading to less accurate selection performance. Finally, none of the existing models on strategy selection considers the influence of different presentation conditions on strategy use. The present work suggests that specifically the availability of an external representation of the task facilitates encoding processes which in turn are crucial for strategy selection mechanisms.

Gaining an accurate understanding of cognitive processes in computational estimation not only yields implications for theoretical models but also has practical ramifications for teaching arithmetic in educational contexts. The production of estimates requires conceptual and procedural knowledge (Bisanz & LeFevre, 2013; LeFevre et al., 1993; Reys et al., 1982; Sowder & Wheeler, 1989). Conceptual knowledge includes the simplification principle which refers to "the understanding that simplifying a problem to an approximate solution is legitimate in some circumstances" (Ganor-Stern, 2016, p. 2), and the proximity principle which refers to "the understanding that an estimate should be as close as possible to the exact answer" (Ganor-Stern, 2016, p. 2). Procedural knowledge refers to "the ways to simplify a problem in order to solve it approximately and the compensation procedures to correct for the distortion produced by the simplification procedures" (Ganor-Stern, 2016, p. 2). It is promising that children as young as third and fourth graders are already quite good in finding approximate sums for addition problems regardless of individual differences in EF, problem characteristics, strategy sets, and presentation forms. This indicates that children as young as third graders already developed the basic understanding of rounding procedures, which is the core of procedural knowledge in computational estimation. In turn, we could not find evidence for the use of compensation procedures in children of this age which is in line with findings of previous studies (LeFevre et al., 1993; Reys et al., 1982; Sowder & Wheeler, 1989). Similarly, the present results imply that the integration of the proximity principle into their estimations (i.e., being specifically accurate in estimating) varies with children's cognitive capacities and other task and problem characteristics that involve cognitive processes. Hence, if the educational goal is that children should adjust their strategy use and come up with estimates that are closest to exact sums, these individual variations should be considered. The results of the present study on the link between updating and children's strategy use imply that children with less efficient cognitive functions need additional support to become proficient estimators. In this regard, interactions with strategy and problem parameters indicate that these children have difficulties to consider unit digits of different operands simultaneously and to execute more complex strategy procedures that require different rounding procedures for multiple operands. Therefore, it might be recommendable to adjust instructions or materials to account for the difficulties of these children. As proposed by other authors (Jitendra et al., 2007; Lemaire et al., 2017), different strategies should be instructed and practiced one at a time in children with fewer cognitive resources. However, later lessons should also focus on the simultaneous consideration of multiple stimuli, as this seems to be a problem for these children. As for the variation of presentation format, the present work implies that a permanent, external representation of the task supports children in selecting the best strategy and executing strategies more accurately. Such visual, external aid could help children with less efficient cognitive processes to obtain performance similar to those of children with more efficient cognitive processes under other presentation, supporting some authors' claim to implement supportive materials to improve performance of students with less efficient cognitive processes in classroom settings (e.g., Elliott, Gathercole, Alloway, Holmes, & Kirkwood, 2010). In educational contexts, this might be implemented by providing these children with tasks in written format or repeating the task when needed. In addition, the present work revealed that already cognitively advantaged children benefit even more from such a permanent, external presentation. Again, this might point to the necessity to specifically strengthen strategic or cognitive processes (e.g., attentional focus) of students with less efficient cognitive processes to support their strategic behavior in educational contexts beyond solely providing them with supportive material.

Furthermore, the present work yields important implications for future studies in strategy research. That is, results showed that the available strategy set influenced children's performance. That is, children spontaneously used the mixed-rounding strategy besides the more systematically investigated rounding-down and rounding-up strategies. Additionally, there was no evidence that children considered the operands' unit digits jointly to correct for distortions produced by simple rounding procedures. Restricting children in their strategy choice was most often explained to increase task difficulty and avoid ceiling effects (e.g., Lemaire & Brun, 2014, 2016; Lemaire & Lecacheur, 2011). However, researchers should be aware of the fact that restricting participants' strategy choice might not simply increase task difficulty but require qualitatively different selection processes compared to these in a more naturalistic setting. Varying problem features like increasing the number of operands and the number of digits of operands or like imposing time-pressure could increase task difficulty and manipulate potential demands on EF in an ecologically meaningful way and could be implemented in future studies to vary task difficulty without the need to restrict participants' strategy choice.

Besides these implications, there are certain limitations of the present studies that should be considered and addressed in future research. First, we addressed the so-called task impurity problem in measuring EF (Rabbitt, 2018) by assessing each component with multiple tasks. Still, we assessed more than one component of EF only in Study 2 while in the other studies we focused on the component of working memory updating. Hence, we cannot rule out the possibility that other EF or general cognitive abilities are involved in children's strategy use. In the studies focusing on working memory updating, we extracted scores of a model with a single factor for updating and the overlapping results on the influence of updating on children's strategy use speak in favor of the robustness of the findings. At the same time, previous research on EF found that performance in each of the three core EF (i.e., updating, shifting, and inhibition) could be decomposed into a contribution of the so-called common EF and a unique contribution of the distinct EF (Miyake & Friedman, 2012). As another example, Fisk and Sharp (2004) found a factor reflecting the efficiency of access to long-term memory in addition to the three mentioned core EF. Consequently, the updating score might contain some variance of a common EF factor or another additional factor underlying individual differences. Going on from this, it would be important to empirically assess the relative contributions of particular EF processes to strategy selection and performance and investigate whether different EF processes account for different experimental effects. In Study 2 we assessed the three EF with multiple tasks. However, due to a lack of reliability we had to exclude the tasks measuring inhibition. Problems regarding the reliability of inhibition scores as individual difference measures are not specific to the present work but are starting to be acknowledged more widely (for a detailed discussion, see also Hedge, Powell, & Sumner, 2018). Also, inconsistencies in finding an inhibition factor have been reported in previous studies (Huizinga, Dolan, & van der Molen, 2006; van der Sluis, de Jong, & van der Leij, 2007; Miyake & Friedmann, 2012). Therefore, future studies should investigate unique contributions of the distinct EF processes in participants' strategy use by including multiple tasks for the various components. As we failed to find significant effects of shifting and working memory on children's strategy use in Study 2, this goal should be pursued while including a larger sample and using factorial analyses to extract scores for each EF to minimize the influence of task-specific variance. Note as well that we examined the three core EF proposed by Miyake and colleagues (2000) in the present work as they are most commonly discussed in the literature. Obviously, this model might not be comprehensive as there are likely other EF that play a role in participants' strategy use. Thus, one could think of many more processes to include in future studies, such as participants' attentional focus (e.g., Wiley & Jarosz, 2012).

Furthermore, there are many other parameters that influence participants' strategy use that have not been addressed in the present work. That is, the context clearly determines the adequacy of an answer and the strategy use therein. For instance, if participants are asked not to produce an estimate but to estimate whether the answer to a problem was larger or smaller than a reference number, in addition to rounding strategies they use a sense of magnitude strategy for certain problems (Ganor-Stern, 2016) which likely involves different cognitive processes. Also, in the present work we did not impose certain situational demands that influence the adequacy of different strategies. As an example, if the task is to answer to the question 'is my money enough for the items in my shopping cart?', the rounding-up strategy or compensation strategies would be more reasonable in this context. In turn, if the task focuses on speed, strategies with less complex procedures would be adequate. Hence, one could think of a future study varying these situational demands as different contexts require different strategies and might give some insight into children's understanding of the usefulness of computational estimation. Finally, the generality of the present findings on cognitive processes in children's strategy use could be tested in a number of other cognitive domains in which participants have been found to use several strategies, such as serial recall and problem-solving (for an overview, see Siegler, 2007).

In sum, future research with further variations in task design could provide a more detailed picture of the cognitive processes underlying participants' strategy use. The present study could shed new light onto some of these cognitive processes.

4.3 Conclusion

As computational estimation is an important skill in everyday life and is connected to other components of mathematical cognition (Siegler & Booth, 2005; Star et al., 2009) it is important to understand potential influential factors. Hence, the aim of the current work was to investigate the cognitive processes that contribute to children's strategy use in computational estimation. For that purpose, in four studies, we investigated the influence of (1) individual differences in EF and (2) varying problem and task characteristics on third and fourth graders' strategy use in computational estimation.

Results showed that children did not spontaneously use compensation when asked to use strategies, implying that associative processes were based on operands' unit digits separately rather than the sum of unit digits. Furthermore, specifically working memory updating seems to contribute to their strategy use. According to interactions with different task variations, updating most likely influences associative processes, long-term memory consolidation and retrieval as well as encoding and calculation processes.

It is promising that children as young as third and fourth graders are quite good in finding approximate sums for addition problems. This is more optimistic than some previous findings on estimation skills in children that age would have implied (Case & Sowder, 1990; Dowker, 1997, 2012; LeFevre et al., 1993). Even though we did not assess children's conceptual understanding of computational estimation, the results indicate that children were in fact able use adequate simplification strategies to obtain an approximate answer. Still, teachers should assure that children understand the meaningfulness and adequacy of simplifying tasks. Hence, before teaching formal rounding procedures, teachers should ensure that students develop the conceptual understanding of computational estimation and develop a quantitative intuition for numbers (Ganor-Stern, 2016; Threadgill-Sowder, 1984). If the educational goal is then to estimate as close and as practical as possible, the present work could shed some light onto cognitive processes that are involved and that can be influenced by certain task and problem parameters.

5 REFERENCES

Adams, J. W., & Hitch, G. J. (1997). Working memory and children's mental addition. *Journal of Experimental Child Psychology*, *67*(1), 21–38. https://doi.org/10.1006/jecp.1997.2397

Ai, J., Yang, J., Zhang, T., Si, J., & Liu, Y. (2017). The effect of central executive load on fourth and sixth graders' use of arithmetic Strategies. *Psychologica Belgica*, *57*(2), 154–172. https://doi.org/10.5334/pb.360

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition*, *44*(1-2), 75–106. https://doi.org/10.1016/0010-0277(92)90051-I

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559. https://doi.org/10.1126/science.1736359

Baddeley, A. (1996). Exploring the Central Executive. *The Quarterly Journal of Experimental Psychology Section a*, *49*(1), 5–28. https://doi.org/10.1080/713755608

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation. The psychology of learning and motivation: Advances in research and theory* (pp. 47–89). Academic Press.

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General*, *133*(1), 83–100. https://doi.org/10.1037/0096-3445.133.1.83

Barrouillet, P., & Lépine, R. (2005). Working memory and children's use of retrieval to solve addition problems. *Journal of Experimental Child Psychology*, *91*(3), 183–204. https://doi.org/10.1016/j.jecp.2005.03.002

Bisanz, J., & LeFevre, J.-A. (2013). Strategic and nonstrategic processing in the development of mathematical cognition. In D. F. Bjorklund (Ed.), *Children's Strategies: Contemporary Views of Cognitive Development* (pp. 213–243). Psychology Press.

Bull, R., & Lee, K. (2014). Executive functioning and mathematics achievement. *Child Development Perspectives*, *8*(1), 36–41. https://doi.org/10.1111/cdep.12059

Campbell, J. I. D. (1992). In defense of the encoding-complex approach: Reply to McCloskey, Macaruso, & Whetstone. In J. I. D. Campbell (Ed.), *The nature and origin of mathematical skills* (pp. 539–556). Elsevier.

Campbell, J. I. D. (1994). Architectures for numerical cognition. *Cognition*, *53*(1), 1–44. https://doi.org/10.1016/0010-0277(94)90075-2

Case, R., & Sowder, J. T. (1990). The development of computational estimation: A neo-piagetian analysis. *Cognition and Instruction*, *7*(2), 79–104. https://doi.org/10.1207/s1532690xci0702_1

Cragg, L., & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. *Trends in Neuroscience and Education*, *3*(2), 63–68. https://doi.org/10.1016/j.tine.2013.12.001

DeStefano, D., & LeFevre, J.-A. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology*, *16*(3), 353–386. https://doi.org/10.1080/09541440244000328

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

Dowker, A. (1997). Young children's addition estimates. *Mathematical Cognition*, *3*(2), 140–153. https://doi.org/10.1080/135467997387452

Dowker, A. (2012). Young children’s estimates for addition: The zone of partial knowledge and understanding. In A. J. Baroody & A. Dowker (Eds.), *Studies in mathematical thinking and learning. The development of arithmetic concepts and skills: Constructing adaptive expertise* (pp. 243–266). Routledge.

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, *14*(4), 172–179. https://doi.org/10.1016/j.tics.2010.01.004

Duverne, S., Lemaire, P., & Vandierendonck, A. (2008). Do working-memory executive components mediate the effects of age on strategy selection or on strategy execution? Insights from arithmetic problem solving. *Psychological Research*, *72*(1), 27–38. https://doi.org/10.1007/s00426-006-0071-5

Ecker, U. K. H., Lewandowsky, S., Oberauer, K., & Chee, A. E. H. (2010). The components of working memory updating: An experimental decomposition and individual differences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*(1), 170–189. https://doi.org/10.1037/a0017891

Elliott, J. G., Gathercole, S. E., Alloway, T. P., Holmes, J., & Kirkwood, H. (2010). An evaluation of a classroom-based intervention to help overcome working memory difficulties and improve long-term academic achievement. *Journal of Cognitive Education and Psychology*, *9*(3), 227–250. https://doi.org/10.1891/1945-8959.9.3.227

Fisk, J. E., & Sharp, C. A. (2004). Age-related impairment in executive functioning: Updating, inhibition, shifting, and access. *Journal of Clinical and Experimental Neuropsychology*, *26*(7), 874–890. https://doi.org/10.1080/13803390490510680

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186–204. https://doi.org/10.1016/j.cortex.2016.04.023

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

Fürst, A. J., & Hitch, G. J. (2000). Separate roles for executive and phonological components of working memory in mental arithmetic. *Memory & Cognition*, *28*(5), 774–782. https://doi.org/10.3758/bf03198412

Ganor-Stern, D. (2016). Solving math problems approximately: A developmental perspective. *PloS One*, *11*(5), 1–16. https://doi.org/10.1371/journal.pone.0155515

Geary, D. C. (1994). *Children's mathematical development: Research and practical applications*. American Psychological Association. https://doi.org/10.1037/10163-000

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*(3), 1166–1186. https://doi.org/10.3758/s13428-017-0935-1

Hodzik, S., & Lemaire, P. (2011). Inhibition and shifting capacities mediate adults' age-related differences in strategy selection and repertoire. *Acta Psychologica*, *137*(3), 335–344. https://doi.org/10.1016/j.actpsy.2011.04.002

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, *44*(11), 2017–2036. https://doi.org/10.1016/j.neuropsychologia.2006.01.010

Imbo, I., & Vandierendonck, A. (2007a). The development of strategy use in elementary school children: Working memory and individual differences. *Journal of Experimental Child Psychology*, *96*(4), 284–309. https://doi.org/10.1016/j.jecp.2006.09.001

Imbo, I., & Vandierendonck, A. (2007b). The role of phonological and executive working memory resources in simple arithmetic strategies. *European Journal of Cognitive Psychology*, *19*(6), 910–933. https://doi.org/10.1080/09541440601051571

Jitendra, A. K., Griffin, C. C., Haria, P., Leh, J., Adams, A., & Kaduvettoor, A. (2007). A comparison of single and multiple strategy instruction on third-grade students' mathematical problem solving. *Journal of Educational Psychology*, *99*(1), 115–127. https://doi.org/10.1037/0022-0663.99.1.115

Kalaman, D. A., & LeFevre, J.-A. (2007). Working memory demands of exact and approximate addition. *European Journal of Cognitive Psychology*, *19*(2), 187–212. https://doi.org/10.1080/09541440600713445

Karbach, J., & Kray, J. (2016). Executive functions. In T. Strobach & J. Karbach (Eds.), *Cognitive training: An overview of features and applications* (Vol. 501, pp. 93–103). Cham: Springer. https://doi.org/10.1007/978-3-319-42662-4_9

Klingner, J., Tversky, B., & Hanrahan, P. (2011). Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks. *Psychophysiology*, *48*(3), 323–332. https://doi.org/10.1111/j.1469-8986.2010.01069.x

Lechuga, M. T., Pelegrina, S., Pelaez, J. L., Martin-Puga, M. E., & Justicia, M. J. (2016). Working memory updating as a predictor of academic attainment. *Educational Psychology*, *36*(4), 675–690. https://doi.org/10.1080/01443410.2014.950193

LeFevre, J.-A., Greenham, S. L., & Waheed, N. (1993). The development of procedural and conceptual knowledge in computational estimation. *Cognition and Instruction*, *11*(2), 95–132. https://doi.org/10.1207/s1532690xci1102_1

LeFevre, J.-A., Lei, Q., Smith-Chant, B. L., & Mullins, D. B. (2001). Multiplication by eye and by ear for Chinese-speaking and English-speaking adults. *Canadian Journal of Experimental Psychology*, *55*(4), 277–284. https://doi.org/10.1037/h0087374

Lemaire, P. (2018). *Cognitive development from a strategy perspective: A Festschrift for Robert Siegler*. *Psychology Press Festschrift series*. London: Routledge. https://doi.org/10.4324/9781315200446

Lemaire, P., Arnaud, L., & Lecacheur, M. (2004). Adults' age-related differences in adaptivity of strategy choices: Evidence from computational estimation. *Psychology and Aging*, *19*(3), 467–481. https://doi.org/10.1037/0882-7974.19.3.467

Lemaire, P., & Brun, F. (2014). Effects of strategy sequences and response-stimulus intervals on children's strategy selection and strategy execution: A study in computational estimation. *Psychological Research*, *78*(4), 506–519. https://doi.org/10.1007/s00426-013-0501-0

Lemaire, P., & Brun, F. (2016). Effects of problem presentation durations on arithmetic strategies: A study in young and older adults. *Journal of Cognitive Psychology*, *28*(8), 909–922. https://doi.org/10.1080/20445911.2016.1215318

Lemaire, P., & Lecacheur, M. (2002). Children's strategies in computational estimation. *Journal of Experimental Child Psychology*, *82*(4), 281–304. https://doi.org/10.1016/S0022-0965(02)00107-8

Lemaire, P., & Lecacheur, M. (2010). Strategy switch costs in arithmetic problem solving. *Memory & Cognition*, *38*(3), 322–332. https://doi.org/10.3758/MC.38.3.322

Lemaire, P., & Lecacheur, M. (2011). Age-related changes in children's executive functions and strategy selection: A study in computational estimation. *Cognitive Development.* https://doi.org/10.1016/j.cogdev.2011.01.002

Lemaire, P., Lecacheur, M., & Farioli, F. (2000). Children's strategy use in computational estimation. *Canadian Journal of Experimental Psychology*, *54*(2), 141–148. https://doi.org/10.1037/h0087336

Lemaire, P., Luwel, K., & Brun, F. (2017). Does the Number of Available Strategies Change How Children Perform Cognitive Tasks? Insights from Arithmetic. *Journal of Educational and Developmental Psychology*, *7*(2), 43. https://doi.org/10.5539/jedp.v7n2p43

Lemaire, P., & Reder, L. (1999). What effects strategy selection in arithmetic? The example of parity and five effects on product verification. *Memory & Cognition*, *27*(2), 364–382. https://doi.org/10.3758/bf03211420

Logie, R. H., Gilhooly, K. J., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory & Cognition*, *22*(4), 395–410. https://doi.org/10.3758/BF03200866

Lovett, M. C., & Anderson, J. R. (1996). History of success and current context in problem solving. *Cognitive Psychology*, *31*(2), 168–217. https://doi.org/10.1006/cogp.1996.0016

Luwel, K., Schillemans, V., Onghena, P., & Verschaffel, L. (2009). Does switching between strategies within the same task involve a cost? *British Journal of Psychology*, *100*(4), 753–771. https://doi.org/10.1348/000712609X402801

Mayer, R. E. (Ed.) (2014). *The Cambridge handbook of multimedia learning*. Cambridge University Press.

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

Paivio, A., & Lacey, K. H. (2007). *Mind and its evolution: A dual coding theoretical approach*. Psychology Press.

Payne, J. W., Johnson, E. J., & Bettman, J. R. (1993). *The adaptive decision maker*. Cambridge University Press.

Rabbitt, P. (Ed.) (2018). *Methodology of frontal and executive function*. Routledge.

Reys, R. E., Rybolt, J. F., Bestgen, B. J., & Wyatt, J. W. (1982). Processes used by good computational estimators. *Journal for Research in Mathematics Education*, *13*(3), 183. https://doi.org/10.2307/748555

Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, *135*(2), 207–236. https://doi.org/10.1037/0096-3445.135.2.207

Seethaler, P. M., & Fuchs, L. S. (2006). The cognitive correlates of computational estimation skill among third-grade students. *Learning Disabilities Research & Practice*, *21*(4), 233–243. https://doi.org/10.1111/j.1540-5826.2006.00220.x

Shrager, J., & Siegler, R. S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science*, *9*(5), 405–410. https://doi.org/10.1111/1467-9280.00076

Siegler, R. S. (2007). Cognitive variability. *Developmental Science*, *10*(1), 104–109. https://doi.org/10.1111/j.1467-7687.2007.00571.x

Siegler, R. S., & Booth, J. L. (2005). Development of numerical estimation. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 197–212). Psychology Press.

Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 229–287). Lawrence Erlbaum Associates.

Sowder, J. T., & Wheeler, M. M. (1989). The development of concepts and strategies used in computational estimation. *Journal for Research in Mathematics Education*, *20*(2), 130–146. https://doi.org/10.2307/749278

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology (2006)*, *59*(4), 745–759. https://doi.org/10.1080/17470210500162854

Star, J. R., Rittle-Johnson, B., Lynch, K., & Perova, N. (2009). The role of prior knowledge in the development of strategy flexibility: the case of computational estimation. *ZDM - the International Journal on Mathematics*, *41*(5), 569–579. https://doi.org/10.1007/s11858-009-0181-9

Threadgill-Sowder, J. (1984). Computational estimation procedures of school children. *The Journal of Educational Research*, *77*(6), 332–336. https://doi.org/10.1080/00220671.1984.10885551

Uittenhove, K., & Lemaire, P. (2013). Strategy sequential difficulty effects vary with working-memory and response-stimulus-intervals: A study in arithmetic. *Acta Psychologica*, *143*(1), 113–118. https://doi.org/10.1016/j.actpsy.2013.02.007

Unsworth, N. (2019). Individual differences in long-term memory. *Psychological Bulletin*, *145*(1), 79–139. https://doi.org/10.1037/bul0000176

Unsworth, N. (2016). Working memory capacity and recall from long-term memory: Examining the influences of encoding strategies, study time allocation, search efficiency, and monitoring abilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(1), 50–61. https://doi.org/10.1037/xlm0000148

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2013). Working memory capacity and retrieval from long-term memory: The role of controlled search. *Memory & Cognition*, *41*(2), 242–254. https://doi.org/10.3758/s13421-012-0261-x

Van der Sluis, S., Jong, P. F. de, & van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence*, *35*(5), 427–449. https://doi.org/10.1016/j.intell.2006.09.001

Van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., & Leseman, P. P. M. (2012). The development of executive functions and early mathematics: A dynamic relationship. *The British Journal of Educational Psychology*, *82*(1), 100–119. https://doi.org/10.1111/j.2044-8279.2011.02035.x

Wiley, J., & Jarosz, A. F. (2012). Working Memory Capacity, Attentional Focus, and Problem Solving. *Current Directions in Psychological Science*, *21*(4), 258–262. https://doi.org/10.1177/0963721412447622

Xu, C., Wells, E., LeFevre, J.-A., & Imbo, I. (2014). Strategic flexibility in computational estimation for Chinese- and Canadian-educated adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(5), 1481–1497. https://doi.org/10.1037/a0037346

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

6 ORIGINAL MANUSCRIPTS

6.1 Study 1: Effects of Working Memory Updating on Strategy Use

Hammerstein, S., Poloczek, S., Lösche, P., Lemaire, P., & Büttner, G. (2019). Effects of working memory updating on children's arithmetic performance and strategy use: A study in computational estimation. *Journal of Experimental Child Psychology, 184,* 174-191.

Journal of Experimental Child Psychology 184 (2019) 174–191

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

Effects of working memory updating on children's arithmetic performance and strategy use: A study in computational estimation

Svenja Hammerstein [a,b,*], Sebastian Poloczek [a,b,c], Patrick Lösche [a,b], Patrick Lemaire [d], Gerhard Büttner [a,b]

[a] Goethe Universität Frankfurt am Main, 60629 Frankfurt am Main, Germany
[b] Centre for Individual Development and Adaptive Education of Children at Risk (IDeA), 60486 Frankfurt am Main, Germany
[c] University of Bristol, Bristol BS8 1TH, UK
[d] Aix-Marseille Université, 13007 Marseille, France

ARTICLE INFO

Article history:
Received 6 February 2019
Revised 3 April 2019

Keywords:
Arithmetic
Strategy use
Development
Working memory
Executive functions
Computational estimation

ABSTRACT

The current study investigated how children's working memory updating processes influence arithmetic performance and strategy use. Large samples of third and fourth graders were asked to find estimates of two-digit addition problems (e.g., 42 + 76). On each problem, children could choose between the rounding-down strategy (i.e., rounding both operands down to the closest decades) or the rounding-up strategy (i.e., rounding both operands up to the closest decades). Four tasks were used to assess updating. Analyses of strategy use revealed that children with more efficient updating showed higher levels of (a) strategy flexibility (i.e., they were less likely to use a single strategy on all or nearly all problems within a test block), (b) strategy adaptivity (i.e., they selected the better strategy overall more often and were more adaptive specifically on homogeneous and rounding-up problems), and (c) strategy performance (i.e., they tended to execute strategies more quickly, especially on homogeneous and larger problems). Finally, updating exerted a more important role for problem type effects in younger children than in older children. These findings have important implications for further understanding how working

* Corresponding author at: Goethe Universität Frankfurt am Main, 60629 Frankfurt am Main, Germany.
E-mail address: hammerstein@psych.uni-frankfurt.de (S. Hammerstein).

https://doi.org/10.1016/j.jecp.2019.04.003

memory updating processes influence children's arithmetic performance and age-related differences therein.

Introduction

Previous works have documented processes involved in children's arithmetic strategy use. Thus, we know that both long-term memory and working memory processes influence children's strategy performance. However, it is unknown whether one specific working memory process, namely updating, influences children's performance and age-related changes. In the current study, we tested the hypothesis that updating processes of working memory play a crucial role in children's arithmetic strategy use. To achieve this end, we adopted a strategy approach that enabled us to examine whether individual differences in children's updating influence which strategy children use and how they execute available strategies while they accomplish a computational estimation task (i.e., finding approximate answers to arithmetic problems). We pursued this goal in the domain of computational estimation because previous works (a) already documented age-related changes in children's performance while trying to find approximate answers to arithmetic problems and (b) showed that children use several strategies to accomplish computational estimation tasks. Before outlining the logic of the current work, we briefly review relevant findings on executive control (EC) processes and arithmetic strategy use and on children's strategic development in arithmetic.

Executive control and arithmetic

EC involves higher-order mental operations concerned with the maintenance, manipulation, planning, monitoring, and regulation of other cognitive processes responsible for perception, memory, reasoning, problem solving, language, and action. These processes involve self-regulation, planning, organization, and the ability to initiate, maintain, switch, and stop sequences of complex behavior. Core EC processes include (a) inhibitory control (resisting habits, temptations, or distractions), (b) working memory (mentally holding and processing information), and (c) cognitive flexibility (for overviews, see Diamond, 2013; Miyake et al., 2000). Updating, one working memory process that is the focus of the current study, helps to keep track of relevant and irrelevant task components and is responsible for actively manipulating contents in working memory (Lehto, 1996; Morris & Jones, 1990).

A number of previous works has found that EC processes are crucial in arithmetic performance in general and in arithmetic strategy use in particular. For example, Imbo and Vandierendonck (2007) asked fourth, fifth, and sixth graders to solve simple addition problems such as 8 + 6 either in a single-task condition (solving only arithmetic problems) or in a dual-task condition (solving arithmetic problems and accomplishing a continuous choice reaction time task, taxing the central executive of working memory). Results revealed that children's performance in the dual-task condition dropped significantly, especially in the youngest children (see Ai, Yang, Zhang, Si, & Liu, 2017, for similar results). As another example, Barrouillet and Lépine (2005) found that children with high memory spans tended to use retrieval more often than children with low memory spans. Both of these studies show that working memory processes are crucially involved in children's strategy use. Finally, Lemaire and Lecacheur (2011) found that strategy choices and strategy execution were influenced by two specific EC processes, namely inhibition and flexibility. Children with better inhibition and cognitive flexibility were more likely to select the better strategy on each problem and to efficiently execute each available strategy when asked to find approximate sums for two-digit addition problems. Thus, it can be concluded that EC processes play a role in children's arithmetic strategy use. However, it is unknown whether working memory updating, one specific EC process, is crucially involved in arithmetic strategy use, an issue that we pursued in the current study.

We pursued this issue in the context of computational estimation. Computational estimation consists in giving approximate answers to arithmetic problems without actually calculating the exact answers but simplifying the calculation (e.g., via rounding the operands). Finding an approximate sum for an addition problem involves encoding the original problem, actively manipulating the operands by rounding the summands, holding the rounded numbers in working memory, and adding the rounded operands before providing an answer. In the current study, we hypothesized that updating processes are crucial to execute this series of processes. Indeed, updating consists in "coding incoming information for relevance to the task at hand and then appropriately revising the items held in working memory by replacing old, no longer relevant information with newer, more relevant information" (Miyake et al., 2000, p. 57). Hence, storing and retrieving intermediate results in and from working memory while solving computational estimation should be facilitated in children with more efficient updating processes. In addition, updating should be involved when activating available strategies in working memory and choosing among strategies as a function of problem characteristics by retrieving associations from long-term memory.

Although no studies have yet investigated how working memory updating is involved in participants' strategy use in a computational estimation task, during past years various studies showed that of all EC processes, updating seems to be most strongly linked to arithmetic performance in general. For example, in a study with 11- and 12-year-old children, St Clair-Thompson and Gathercole (2006) identified high unique associations of a factor including updating and working memory span measures with mathematics scores in a principal component analysis. The mathematics score was operationalized by one mental and two written arithmetic tests. Although St Clair-Thompson and Gathercole did not use a computational estimation task to assess children's arithmetic abilities, we expect performance specifically in the mental test to be highly linked to performance in computational estimation because mental manipulation of the task and mental addition are required when finding approximate sums for addition problems. As another example, a meta-analysis by Friso-van den Bos, van der Ven, Kroesbergen, and van Luit (2013) revealed a higher correlation between mathematics and updating than between mathematics and inhibition or shifting. Furthermore, a confirmatory factor analysis in second graders showed that the development of updating was strongly related to mathematical development, whereas inhibition and shifting did not predict arithmetic, when controlling for updating (van der Ven, Kroesbergen, Boom, & Leseman, 2012). Moreover, Lechuga, Pelegrina, Pelaez, Martin-Puga, and Justicia (2016) investigated the relative contributions of working memory updating and intelligence to academic attainment in fourth graders. In hierarchical regression analyses, they found that both variables made a unique contribution to the prediction of mathematical problem solving and arithmetical operations. Importantly, updating accounted for a larger amount of variance in both arithmetic performance variables. Again, the arithmetic scores comprised various math tasks such as simple addition, mental calculation, and number comparison, which are relevant subprocesses within computational estimation.

In sum, previous research found that EC processes crucially influence children's arithmetic performance in the domain of strategies and age-related changes in this performance. This influence occurs via children with the most efficient EC processes using the best strategy more often and executing available strategies more efficiently. One of the important limitations of previous studies, however, is that none investigated the role of one EC process, namely updating. As a consequence, we do not know whether updating is, like other core EC processes, crucial in children's arithmetic strategy use. The possibility that updating may be crucial for arithmetic strategy use has been discussed by several authors (e.g., DeStefano & LeFevre, 2004). However, no empirical data have documented this role. Therefore, the aim of the current study was to investigate the contribution of children's working memory updating processes to their arithmetic performance as well as to their arithmetic strategy use. We adopted a strategy approach because previous studies established the crucial role of strategic aspects of children's performance in arithmetic.

Children's strategic development in arithmetic

An important factor in children's academic performance in general, and in math in particular, is cognitive strategy use. A strategy can be defined as "a procedure or set of procedures for achieving a higher level goal or task" (Lemaire & Reder, 1999, p. 365). Previous research found that children use various strategies while performing arithmetic problem solving as well as many other cognitive tasks. With increasing age, children tend to rely less on easier strategies, select the best strategy for a problem more often, and execute strategies more efficiently (for overviews, see Lemaire, 2017; Siegler, 2007). As previously outlined, an open question is whether these developmental effects can be accounted for by the development of cognitive control processes such as updating. Some previous research indicates that children's use of cognitive strategies is influenced by children's EC processes (Lemaire & Lecacheur, 2011).

Moreover, there is evidence that different problem features, such as problem size and problem type, influence participants' strategies and performance in arithmetic (LeFevre, Greenham, & Waheed, 1993; Lemaire, Lecacheur, & Farioli, 2000; Lemaire & Lecacheur, 2011). For example, participants perform better on problems with smaller magnitudes (e.g., 27 + 12) than on problems with larger magnitudes (e.g., 87 + 62). In addition, they choose the better strategy more often on homogeneous problems (i.e., problems for which unit digits of operands are either both smaller or both larger than 5) than on heterogeneous problems (i.e., problems for which the unit digit is smaller than 5 in one operand and larger than 5 in the other operand). Determining which strategy is better on each problem and executing strategies is known to place demands on EC resources (e.g., Lemaire & Lecacheur, 2011). Here we tested the hypothesis that EC plays a crucial role in arithmetic performance via its effects on strategy use and strategy execution by comparing how children with varying efficiency levels in EC accomplish our computational task. In addition, items were carefully constructed, in order to vary the demands on EC.

Above and beyond documenting empirical relations between EC and children's arithmetic, our findings were expected to contribute to theories of strategies. Several models of strategies have been proposed to account for children's strategy selection as well as their strategic development. These are the adaptive control of thought–rational (ACT-R) model by Lovett and Anderson (1996), the strategy selection learning (SSL) model by Rieskamp and Otto (2006), and the strategy choice and discovery simulation (SCADS) model by Shrager and Siegler (1998). These models have several common core assumptions. For instance, they assume that children select strategies on a trial-by-trial basis. This means that children select the strategy on each trial rather than a priori choosing one strategy to solve all or nearly all problems. In addition, models propose that when children are presented with a problem, they activate relative benefits and costs of available strategies and select the strategy with larger benefits and lower costs. Strategic development is commonly explained as the result of age-related changes in processes from learning experience. Such processes include "associative learning mechanisms" (Shrager & Siegler, 1998, p. 407), "reinforcement learning" (Rieskamp & Otto, 2006, p. 208), and "processing of experience-based information" (Lovett & Anderson, 1996, p. 169). Previous studies documented sources of interindividual differences during development such as inhibition and flexibility (e.g., Lemaire & Lecacheur, 2011). Models of strategic development have not been devised to test the role of these processes and, as a consequence, do not say much about them as an explanation for strategic development. However, as already mentioned, because previous empirical studies established that EC processes are crucial to strategic development, by investigating the role of working memory updating, the current study was expected to shed important light on theories of how strategic changes occur. Furthermore, we argue that EC processes contribute to existing models by explaining several components during strategy selection and execution. For example, to choose strategies on each trial, participants need to switch strategies from one trial to the next (Lemaire & Lecacheur, 2011). To achieve this, they need to inhibit the just executed strategy on a given problem, reactivate the set of available strategies in working memory, and choose among these strategies as a function of problem characteristics by retrieving associations from long-term memory. In addition, participants need to coordinate the set of processes involved in a given strategy during strategy execution, for which EC processes may be crucial.

In sum, previous theoretical and empirical works on arithmetic have shown that participants' performance is influenced by the strategies they use, the type of problems they solve, participants' age, and general processing resources such as working memory capacities and some EC processes (e.g., inhibition, flexibility). It is unknown how one type of EC processes, namely working memory updating, influences children's arithmetic performance and whether updating contributes to strategic processes in arithmetic and to differences in performance as a function of children's age and problem characteristics. We pursued these issues in the current study.

The current study

To examine relations between updating and arithmetic, we asked large samples of third and fourth graders to find approximate sums for two-digit addition problems (e.g., 42 + 76). On each problem, children could choose between two available strategies, namely the rounding-down strategy (i.e., rounding both summands down to the closest decades such as doing 40 + 70 = 110 for 42 + 76) and the rounding-up strategy (i.e., rounding both summands up to the closest decades such as doing 50 + 80 = 130 for 42 + 76). We included two types of problems, namely so-called homogeneous and heterogeneous problems. These two problem types were included following previous findings that children's performance and strategies differ between these types of problems and because the problem type effect changes with age; children are more likely to select the better strategy on homogeneous problems than on heterogeneous problems, and this effect is larger in older children (Lemaire & Brun, 2014; Lemaire & Lecacheur, 2011). Moreover, half of the problems were so-called rounding-down problems (i.e., the use of the rounding-down strategy would lead to better estimates that are closer to the exact sum) and half were rounding-up problems (i.e., rounding up would lead to better estimates) because previous works showed that children select the better strategy more often on rounding-down problems and are faster when using the rounding-down strategy (Lemaire & Brun, 2016). To examine problem size effects, we used a wide range of exact sums (i.e., from 58 to 163).

We decided to investigate third and fourth graders following previous studies showing that important strategic changes occur in children of this age range (e.g., LeFevre et al., 1993; Lemaire & Lecacheur, 2011). Moreover, in German schools, where the data were collected, rounding rules and computational estimation are first introduced in third grade and further practiced in fourth grade. Therefore, if working memory updating is crucial in strategy use and strategic changes, it was expected that it would be most sensitively detected in third and fourth graders.

Most originally, we assessed children's working memory updating functions with four updating tasks. An indicator of updating underlying all four tasks was extracted to obtain a more reliable and valid measure of updating. By combining measures of four different updating tasks, we minimized the influence of task-specific processes to address the so-called task impurity problem within the field of executive functions (Rabbitt, 1997). We tested the hypothesis that children with more efficient updating would choose the better strategy more often and execute strategies more efficiently. Another goal was to examine whether the efficiency of updating moderates effects of item characteristics (i.e., rounding-down/rounding-up problems, homogeneous/heterogeneous problems, and smaller/larger problems). Thus, we tested whether children with more efficient updating would show smaller problem size effects, smaller differences between rounding-up and rounding-down problems, and larger problem type effects.

Method

Participants

A total of 308 children were tested: 158 third graders (90 boys; age in months: M = 114.0, SD = 5.6, range = 102.4–138.7) and 150 fourth graders (80 boys; age in months: M = 126.1, SD = 5.2, range = 116.2–147.0). The study was conducted at the end of the school year. Children were recruited from 24 classes in 11 elementary schools in urban and suburban areas of the state of Hesse, Germany. The study was approved by the local ethics committee. Furthermore, parents provided written informed consent and participants gave their verbal assent.

Materials and procedure

All participants completed four updating tasks and a computational estimation task. Children took part in two sessions, each lasting approximately 45 min with an average of 8.5 days (SD = 5.7) in between sessions. The order of test administration was the same for all participants.

Children were tested in groups of 2–11 individuals (*M* = 7.4, *SD* = 2.3).[1] They solved all tasks on tablet computers. Tasks were programmed with E-Prime 3.0. Instructions were presented verbally over headphones, and additional instructions were given when children did not respond or made errors during the first practice trials. Two experimenters were present for each group to answer children's questions and provide further explanations if needed.

Working memory updating tasks

Working memory updating was assessed with four different tasks to obtain a stable and purer measure of updating. These four tasks were the spatial keep track, day keep track, frog position updating, and color updating tasks.

The *spatial keep track* task (Dirk & Schmiedek, 2016) comprised 16 trials. On each trial, children were presented with either two or three differently colored fictitious creatures (i.e., monsters) at different positions on a 4 by 4 grid. Each of the two difficulty levels (i.e., two vs. three creatures) was presented 8 times. Children were asked to remember the initial positions of the creatures. The creatures disappeared after 3000 ms. Then, arrows matching the creatures' colors were successively presented in the center of the grid for 2500 ms with interstimulus intervals of 500 ms. Each arrow indicated that the creature of the corresponding color should move one block in the arrow's direction. Three arrows for two-creature trials and four arrows for three-creature trials were presented (i.e., three and four updating operations, respectively). Participants were asked to remember the new updated positions of all creatures and to reproduce the final positions on an empty grid at the end of each trial.

An arithmetic keep track task (Dirk & Schmiedek, 2016) was adapted and changed into a *day keep track* task to avoid an arithmetic confound in our measures of updating. Across 18 trials, participants were presented with one, two, or three distinct calendar pages. On each page, the German name for a weekday was written. Children were asked to remember the initial weekdays on the calendar pages. The names of the weekdays appeared for 3000 ms for one-calendar trials and for 4000 ms for two- and three-calendar trials. After displaying weekday names, a series of updating instructions (i.e., German words for *1 forward, 2 forward, 1 backward,* and *2 backward* and corresponding arrows) was presented on empty calendar pages for 3000 ms one at a time. Updating operations were presented with interstimulus intervals of 500 ms. Participants were asked to remember the updated weekdays for each calendar and to reproduce the final weekdays. The weekdays needed to be updated three times for one- and two-calendar trials and three or four times for three-calendar trials. Each of these three difficulty levels (i.e., one, two, or three-calendar trials) was presented six times.

Next, in the *frog position updating* task (LeFevre et al., 2009), participants were presented with movement sequences of three to seven distinct positions of a frog on an array of eight irregularly positioned water lily leaves. Children were then asked to indicate the last two, three, or four positions of the frog. Each of these three difficulty levels (i.e., remembering the last two, three, or four positions) was presented five times. Each position of the frog was displayed for 1750 ms with interstimulus intervals of 250 ms.

Finally, children performed 16 trials of a *color updating* task (Lee, Ng, Bull, Pe, & Ho, 2011). They were visually and verbally presented with sequences of four to eight colors one at a time. Children were asked to remember either the last two or three colors in eight trials each. Stimuli consisted of eight colors with monosyllabic names in German. Each color was presented for 1500 ms before it was covered for 500 ms and followed by the presentation of the next color.

The percentage of correctly remembered items in each of the four tasks was used as a measure of individuals' updating function. If more than half of the trials of a difficulty level were missing for a participant (e.g., missing at random due to program crashes), the subtest score was not calculated. Thus,

[1] Following one of the reviewers' comments, we tested the influence of the size of the group on children's strategy use. Analyses revealed that children in larger groups showed faster estimation latencies and that this effect was somewhat larger in fourth graders. However, including the size of the group as a predictor did not change the current main and important results. No effects of the size of the group were found on children's strategy flexibility or better strategy selection (see also the syntax script in the online supplementary material for the analyses with the size of the group included).

subtest scores were calculated for 249, 264, 262, and 275 children out of 277 children for the spatial keep track, day keep track, frog updating, and color updating tasks, respectively. The internal consistency of the updating tasks, as calculated by Cronbach's alpha for two balanced parcels each, was high (αs = .89, .85, .87, and .85 for the spatial keep track, day keep track, frog position updating, and color updating tasks, respectively). Moreover, the internal consistency of the updating composite score, as calculated by the mean of all *z*-standardized parcels, was high (α = .89).

To verify the factorial structure, we ran an exploratory factor analysis including the *z* scores of the updating subtest parcels. The corresponding scree plot clearly indicated a structure with a single factor only. To obtain a task-nonspecific updating value, the factor scores of a factor analysis with a single factor were used for further analyses. We centered updating scores within grades; thus, a score of zero (0) in the following analyses reflects the updating function of an average child within his or her grade.

Computational estimation task

Task stimuli. Children worked on two sets of 21 two-digit addition problems each. Problems included operands ranging from 21 to 89; exact sums ranged from 58 to 163. One third of the problems were homogeneous problems, and two thirds of the problems were heterogeneous problems. Unit digits both were smaller than 5 (e.g., 52 + 63) in half the homogeneous problems and larger than 5 in the other homogeneous problems (e.g., 38 + 76). The unit digit of the first operand was smaller than 5 in the first operand and larger than 5 in the second operand in half of the heterogeneous problems (e.g., 41 + 27), and the reverse was the case in the other heterogeneous problems (e.g., 68 + 43).

Half the problems were so-called rounding-down problems (e.g., 42 + 61), where rounding both operands down to the nearest decades yielded better estimates (i.e., closer to the exact sum). The other problems were so-called rounding-up problems (e.g., 29 + 68), where rounding both operands up to the nearest decades yielded better estimates. Rounding-down and rounding-up problems were matched on mean exact sums (*M*s = 107.2 and 105.2 for rounding-down and rounding-up problems, respectively), on the number of problems requiring a carry over 100, and on their mean splits between unit digits from the nearest decades when rounding with the better strategy (*M* = 7.2 for both rounding-up and rounding-down problems). Moreover, rounding-down and rounding-up problems were matched on mean percentage deviations when solved with the better and poorer strategies. Thus, mean percentage deviations between correct sums and estimates for rounding-down problems were 7.3 (range = 2.8–13.8) when solved with the rounding-down strategy and 12.8 (range = 7.9–20.7) when solved with the rounding-up strategy. Mean percentage deviations between correct sums and estimates for rounding-up problems were 13.2 (range = 7.3–23.1) when solved with the rounding-down strategy and 7.5 (range = 2.2–12.7) when solved with the rounding-up strategy.

Finally, based on previous findings in arithmetic (for overviews, see Cohen Kadosh & Dowker, 2015; Gilmore, Göbel, & Inglis, 2018), we controlled the following factors. First, no operands had a 0 or 5 as unit digit (e.g., 20 + 63, 25 + 63). Second, sums of unit digits never equaled 10. Third, unit and decade digits were never the same within operands (e.g., 44 + 23). Fourth, unit or decade digits were never the same for both operands in a given problem (e.g., 32 + 62, 49 + 41). Fifth, no problems with reverse order of the operands from another problem were presented (e.g., if 68 + 24 was included, 24 + 68 was not). Sixth, no operand had its closest decade equal to 0, 10, or 100.

Task procedure. Children were told that they should give an approximate answer for each addition problem without actually calculating the exact sum. They were instructed to use one of two rounding strategies, either the rounding-down strategy (i.e., rounding both operands down to the nearest decades) or the rounding-up strategy (i.e., rounding both operands up to the nearest decades). The mixed-rounding strategy (i.e., rounding one operand up and the other operand down to the nearest decades) was not allowed. They were asked to use the better strategy for each addition problem. The better strategy was defined as the strategy that yielded the closest estimate to the exact sum. They were also asked to be as fast as possible.

Children were asked to indicate their answer to each problem on a touchscreen-based numpad. Strategies were not assessed directly by asking children which strategy they had used on each problem. Rather, we inferred which strategy was used on each problem from given responses (e.g., a child was considered to have used the rounding-down strategy after providing 90 as a sum estimate for 41 + 58 or the rounding-up strategy after providing 110 as the estimate; and any other estimate was coded as "other"). The numpad consisted of keys with rounded numbers from 10 to 200 to prevent children from typing in exact sums or estimates with 5 as a unit digit. Prior to the computational estimation task, children were asked to enter eight verbally presented two-digit numbers (e.g., 60) on the numpad to ensure that they were familiar with the input modality. Then, children practiced on six training problems. They were provided with feedback on each practice problem regarding whether they had chosen one of the two allowed strategies to ensure that they were familiar with executing the rounding strategies. Although some children initially used the mixed-rounding strategy, all individuals had no difficulties with the procedure or with either rounding down or rounding up. Finally, problems were presented with varying response stimulus intervals (RSIs)[2] from 600 to 2000 ms in random order.

Data processing

An entire test block of a participant (i.e., one set of 21 trials) was excluded from all further analyses if on at least one third of trials values were missing or responses did not match the result of one of the two available rounding strategies. Thus, both test blocks of 21 children and either the first or second test block of 33 children were excluded. In addition, the first trial of each test block was not considered in analyses. Moreover, trials of the computational estimation task were excluded from further analyses regarding their estimation times if the given answer did not match the result of one of the two available rounding strategies (thus, 558 trials were excluded).

To omit trials with extreme individual response times, z–standardized values of the estimation times were calculated within participants. Trials were excluded from estimation time analyses if the values exceeded 2.5 (thus, 299 trials were excluded) because they most likely resulted from children's transient distraction from the task. To handle data of children with an extremely slow response pattern, we calculated z–standardized values of children's mean estimation times within grades. Thus, 10 children were excluded from all further analyses because their z values exceeded 2.5, resulting in a sample size of 277 children.

Children's strategy flexibility was analyzed for these 277 children. Because we decided to focus mainly on children's flexible strategy use, we included only flexible test blocks in analyses of children's better strategy selection and estimation latencies. Thus, the latter analyses were run for a subsample of 226 children with data on at least one flexible test block.

Statistical method

All analyses were run using RStudio. Children's better strategy selection and strategy estimation latencies were examined with cross-classified multilevel models to allow for and to examine (a) between-participant variance and (b) between-item variance. The data have a two-level cross-classified structure, with trials nested within 42 items as well as within 226 participants. Only trials with valid strategy use (i.e., rounding down or rounding up) were considered.

Children's dichotomous strategy selection variables were analyzed with logit models for binary responses. To be able to interpret effects in terms of probabilities that $y = 1$ (e.g., that children used the better strategy on a given problem), we calculated the predicted probabilities of children's strategy selection for the different values of predictors and report them in the following as percentages. We calculated the predicted probabilities as population-averaged probabilities by averaging over different simulated Level 2 random intercepts. Thus, percentages in Results can be interpreted as predicted probabilities that are true for the average of children and items with a certain predictor combination (see also Steele, 2009, for the calculation of predicted probabilities in logit models).

[2] The variation of RSI is not the scope of the current article. Note however, that including RSI as a factor did not change the presented results.

Results

Effects of working memory updating on strategy selection

Strategy variability

To examine children's variability in their strategy use, we calculated the percentage of the rounding-down strategy use for each participant and test block. A score of .60 indicates that the child used the rounding-down strategy on 60% of the problems within the test block. The distribution of these scores revealed a trimodal structure (see the vertical histogram in Fig. 1). This multimodal structure points to the existence of subgroups within the current data. The first subgroup consists of children in the center of the distribution approaching the problems of a test block in a flexible manner and adjusting strategies on a problem-by-problem basis; the further two subgroups share an inflexible approach and solve all or nearly all problems with the same strategy, either rounding down (i.e., high values of percentage use of the rounding-down strategy) or rounding up (i.e., low values of percentage use of the rounding-down strategy). To take the two approaches into consideration, we created a dichotomous inflexibility variable. Because the distribution of the percentage use of the rounding-down strategy revealed minima at about 90% and 10%, a test block was classified as being approached in an inflexible manner when 90% or more of problems were solved with either the rounding-down or rounding-up strategy.

Most third graders ($n = 75$) and fourth graders ($n = 123$) solved problems within both test blocks with a flexible approach. Thus, the current study replicates previous findings on within-participant strategy variability (Siegler, 2007). However, 43 third graders and 8 fourth graders used an inflexible approach on all problems. The remaining 28 children solved one test block with a flexible approach and solved the other test block with an inflexible approach. Because 60 first test blocks and 65 second test blocks were classified as inflexible, inflexibility did not merely stem from effects of task familiarization (i.e., children would show strategy inflexibility more frequently within earlier test blocks) or motivational decline and fatigue (i.e., children would show inflexibility more frequently within later test blocks).

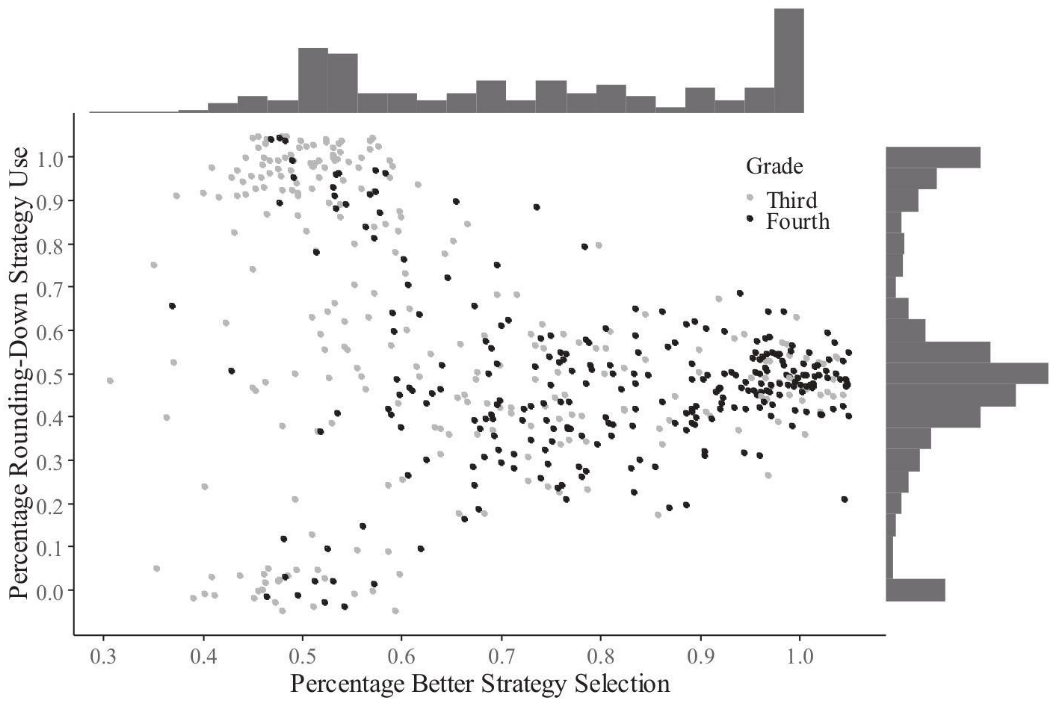


Fig. 1. Relations between children's percentage rounding-down strategy use and percentage better strategy selection. Note that jitter was added in the scatterplot to make overlapping data points visible. This resulted in some data points being displayed beyond 0.0 or 1.0.

Note that in 74% of inflexible test blocks, the most frequently used strategy was the rounding-down strategy (see Fig. 1). However, in the other 26% of inflexible test blocks, the dominant strategy was the more complex rounding-up strategy. This points to the fact that inflexibility did not merely stem from children choosing the easier strategy on each problem because it is associated with less costs. Rather, the avoidance of cognitive costs involved in the process of strategy selection and switching between strategies fits the data.

To test the effects of grade (i.e., third vs. fourth) and working memory updating on children's flexibility on each test block, a logistic multilevel model was fitted to the dataset. Results revealed that it was more likely that a test block was approached in an inflexible manner by a third grader than by a fourth grader (39% vs. 9%; $\beta = -2.86$, 95% confidence interval (CI) [−5.19, −0.53], $p = .02$), and children with less efficient working memory updating tended to approach a test block in an inflexible manner more often than children with more efficient updating (average predicted probability (PP) for children with a standardized updating score of +1 *SD*, $PP_{+1SD} = 27\%$ vs. $PP_{-1SD} = 32\%$; $\beta = -0.95$, 95% CI [−2.05, 0.15], $p = .09$).

Focus on test blocks with a flexible approach

Fig. 1 reveals qualitatively different subgroups. Very high values in the percentage use of a dominant strategy (i.e., classification as inflexible) are (a) likely linked to an approach in which children do not select strategies on a trial-by-trial basis and (b) necessarily tied to a better strategy selection at about 50%, whereas in flexible test blocks the adaptivity generally is clearly higher. Because analyses regarding children's strategy use, therefore, would be distorted when including test blocks that were approached in an inflexible way, we included only flexible test blocks with a dominant strategy use of less than 90% in analyses of children's better strategy selection and estimation latencies. Thus, the following analyses were run for a subsample of 226 children (i.e., 94 third graders and 132 fourth graders) with data on at least one flexible test block.

Better strategy selection

The next analysis aimed at determining whether better strategy selection varied with participant and problem characteristics. Better strategy selection was coded 1 if children selected the better strategy on a problem (i.e., using the rounding-down strategy on rounding-down problems or the rounding-up strategy on rounding-up problems) and 0 otherwise. A logistic cross-classified multilevel model was fitted to the dataset to test the effects of grade (third vs. fourth), working memory updating, problem type (homogeneous vs. heterogeneous), problem size (size of the exact sum), and rounding type (rounding down vs. rounding up).

Fourth graders were more likely to select the better strategy than third graders (PP = 86% vs. 72%; $\beta = 0.65$, 95% CI [0.45, 0.84], $p < .001$), and children with more efficient updating were more likely to use the better strategy than children with less efficient updating ($PP_{+1SD} = 84\%$ vs. $PP_{-1SD} = 74\%$; $\beta = 0.49$, 95% CI [0.29, 0.68], $p < .001$) (see also Fig. 2A). Interestingly, it was equally likely for a third grader with updating 1 *SD* above average to choose the better strategy (PP = 78%) as for a fourth grader with updating 1 *SD* below average (PP = 82%). The Grade × Updating interaction was not significant ($\beta = 0.06$, 95% CI [−0.13, 0.26], $p = .52$). Moreover, children were more likely to use the better strategy on homogeneous problems than on heterogeneous problems (PP = 85% vs. 72%; $\beta = -0.64$, 95% CI [−0.79, −0.49], $p < .001$). Importantly, the Grade × Problem Type interaction ($\beta = -0.24$, 95% CI [−0.32, −0.16], $p < .001$) and the Updating × Problem Type interaction ($\beta = -0.12$, 95% CI [−0.20, −0.03], $p = .006$)[3] were significant. This occurred because the effect of heterogeneous versus homoge-

[3] Note that the change of predicted probabilities with different problem types is numerically comparable in children with less efficient updating ($PP_{homogeneous} = 81\%$ vs. $PP_{heterogeneous} = 67\%$) and in children with more efficient updating ($PP_{homogeneous} = 91\%$ vs. $PP_{heterogeneous} = 78\%$); however, the significant interaction effect reveals that the problem type effect is larger in children with more efficient updating. This occurs because the logit model accounts for the fact that differences of predicted probabilities on the restricted scale from 0% to 100% are less likely in the upper and lower ranges close to 0% and 100%. Thus, the model considers that changes in probabilities are located in different ranges on the restricted probability continuum from 0 to 1 and does not solely compare numerical changes (see also Jaeger, 2008).

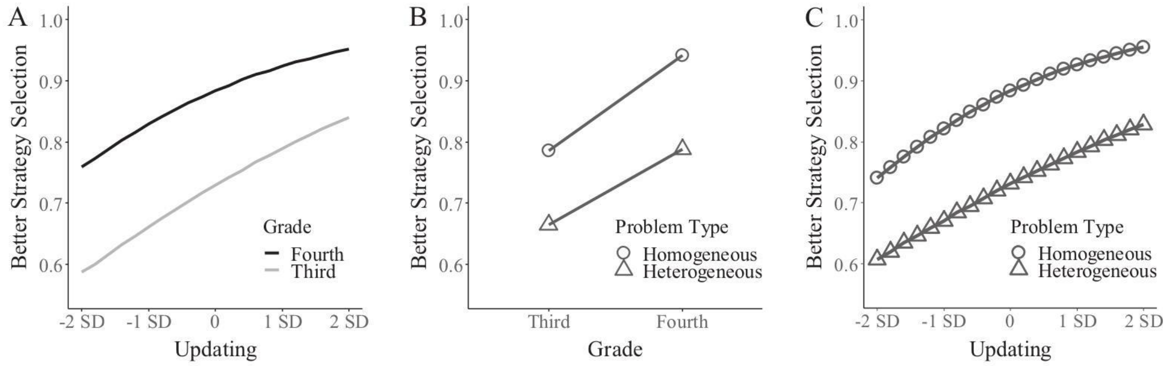


Fig. 2. (A) Influence of updating on third and fourth graders' predicted better strategy selection. (B and C) Influence of grade (B) and updating (C) on predicted better strategy selection while solving heterogeneous or homogeneous problems.

neous problems was larger in older children than in younger children and was larger in children with more efficient updating (see Fig. 2B and C and Table 1). The Grade × Updating × Problem Type interaction (β = −0.02, 95% CI [−0.11, 0.06], p = .64) was not significant.

The main effect of rounding type was not significant (β = −0.03, 95% CI [−0.16, 0.11], p = .59), but the interaction effects Grade × Rounding Type (β = 0.13, 95% CI [0.06, 0.19], p < .001) and Updating × Rounding Type (β = 0.07, 95% CI [0.01, 0.14], p = .04) were significant. This occurred because only younger children were more likely to select the better strategy on rounding-down problems than on rounding-up problems, whereas older children chose the better strategy on rounding-down and rounding-up problems equally often (see Fig. 3A and Table 2). Similarly, the effects of type of rounding problems on better strategy selection were larger in children with less efficient updating than in children with more efficient updating (see Fig. 3B and Table 2). The three-way interaction of Grade × Updating × Rounding Type (β = −0.01, 95% CI [−0.07, 0.06], p = .87) was not significant.

All main and interaction effects including problem size were nonsignificant, with effect estimates close to 0 and narrow confidence intervals: problem size (β = −0.00, 95% CI [−0.05, 0.04], p = .90), Grade × Problem Size (β = −0.01, 95% CI [−0.03, 0.02], p = .59), Updating × Problem Size (β = 0.00, 95% CI [−0.02, 0.03], p = .76), and Grade × Updating × Problem Size (β = 0.00, 95% CI [−0.02, 0.03], p = .88).

Table 1
Predicted probabilities (%) of better strategy selection in third and fourth graders with less versus more efficient updating while solving homogeneous or heterogeneous problems.

		Grade		
	Problem type	Third	Fourth	Mean
Less efficient updating (−1 *SD*)	Homogeneous	71	90	81
	Heterogeneous	61	73	67
	Mean	66	82	74
More efficient updating (+1 *SD*)	Homogeneous	85	97	91
	Heterogeneous	72	84	78
	Mean	78	90	84
Mean	Homogeneous	79	94	85
	Heterogeneous	67	79	72
	Mean	72	86	79

Note. Unlike in an analysis of variance, children were not split into subsamples with less or more efficient updating, but the effect of updating was modeled continuously. For illustration, predicted probabilities are shown for children with updating processes 1 *SD* above average and 1 *SD* below average.

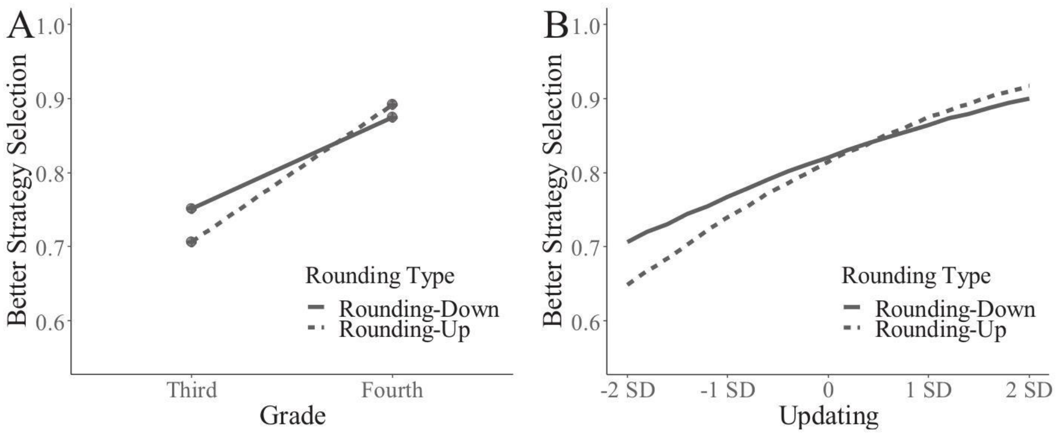


Fig. 3. Influence of grade (A) and updating (B) on children's predicted better strategy selection while solving rounding-down or rounding-up problems.

Table 2
Predicted probabilities (%) of better strategy selection in third and fourth graders with less versus more efficient updating while solving rounding-down or rounding-up problems.

		Grade		
	Rounding type	Third	Fourth	Mean
Less efficient updating (−1 *SD*)	Rounding down	70	81	76
	Rounding up	62	82	72
	Mean	66	82	74
More efficient updating (+1 *SD*)	Rounding down	79	89	84
	Rounding up	77	91	84
	Mean	78	90	84
Mean	Rounding down	75	86	80
	Rounding up	70	87	78
	Mean	72	86	79

Note. Unlike in an analysis of variance, children were not split into subsamples with less or more efficient updating, but the effect of updating was modeled continuously. For illustration, predicted probabilities are shown for children with updating processes 1 *SD* above average and 1 *SD* below average.

Effects of working memory updating on estimation latencies

A cross-classified multilevel model was fitted to the dataset to test the effects of grade (third vs. fourth), working memory updating, problem size (size of the exact sum), problem type (homogeneous vs. heterogeneous), and the strategy used (rounding down vs. rounding up) on children's estimation latencies on each problem. The influence of the predictors within the linear multilevel model can be calculated the same way as in multiple regressions by multiplying beta coefficients with the corresponding predictor values. Thus, estimation latencies can be derived by simple addition of beta coefficients, with all dummy variables being coded −1/+1 and a change of 1 representing a change of 1 *SD* for updating and a change of 10 units for the exact sum. The reported influences on estimation latencies are true for an average child and an average item.

Because the fixed intercept was $\beta = 12.68$, 95% CI [12.10, 13.26], children took on average 12.7 s to estimate a problem and fourth graders were faster than third graders (12.1 vs. 13.2 s; $\beta = -0.55$, 95% CI [−1.06, −0.05], $p = .04$). Moreover, children were faster on homogeneous problems than on heterogeneous problems (11.8 vs. 13.6 s; $\beta = 0.90$, 95% CI [0.58, 1.22], $p < .001$), but this effect was qualified by both a Grade × Problem Type interaction ($\beta = 0.31$, 95% CI [0.15, 0.47], $p < .001$) and an Updating × Problem Type interaction ($\beta = 0.21$, 95% CI [0.04, 0.38], $p = .01$). This occurred because differences in estimation times between heterogeneous and homogeneous problems were larger for

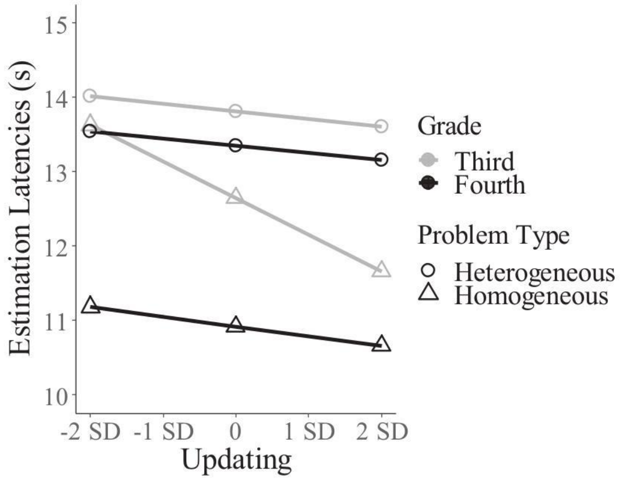


Fig. 4. Influence of updating on children's estimation times for third and fourth graders while solving heterogeneous or homogeneous problems.

older children and for children with more efficient updating than for other children. Indeed, fourth graders were on average 2.4 s faster on homogeneous problems than on heterogeneous problems, whereas third graders were only 1.1 s faster. Similarly, children with updating 1 *SD* above average were 2.2 s faster on homogeneous trials than on heterogeneous trials, whereas children with updating 1 *SD* below average were only 1.4 s faster. Interestingly, the Grade × Updating × Problem Type interaction was significant ($\beta = -0.18$, 95% CI [−0.34, −0.01], $p = .04$). This occurred because the influence of updating functions on the problem type effect was larger in third graders than in fourth graders (see Fig. 4). Indeed, in third grade, children with updating 1 *SD* below average were on average only 0.4 s faster on homogeneous problems than on heterogeneous problems (corresponding difference in latencies was 1.9 s in children with updating 1 *SD* above average). In contrast, these differences were almost the same for fourth graders with updating 1 *SD* below average (2.4 s) and above average (2.5 s). Thus, more efficient updating processes specifically facilitated third graders' performance on homogeneous problems.

Furthermore, children were slower on larger addition problems (on average + 0.34 s for a 10-unit increase of the exact sum; $\beta = 0.34$, 95% CI [0.24, 0.45], $p < .001$). Importantly, this problem size effect interacted with children's updating processes ($\beta = -0.08$, 95% CI [−0.14, −0.03], $p = .003$). This occurred because children with more efficient updating were less influenced by problem size. Thus, children with updating 1 *SD* above average would slow down by only 0.26 s for a 10-unit increase of the exact sum, whereas children with updating 1 *SD* below average would slow down by 0.42 s. In addition, children were faster when using the rounding-down strategy than when using the rounding-up strategy (12.2 vs. 13.2 s; $\beta = 0.49$, 95% CI [0.19, 0.80], $p = .003$).

No other effects were significant: updating ($\beta = -0.43$, 95% CI [−0.95, 0.09], $p = .10$), Grade × Updating ($\beta = 0.20$, 95% CI [−0.31, 0.72], $p = .45$), Grade × Problem Size ($\beta = -0.03$, 95% CI [−0.09, 0.02], $p = .24$), Grade × Updating × Problem Size ($\beta = 0.03$, 95% CI [−0.02, 0.08], $p = .31$), Grade × Strategy Used ($\beta = -0.01$, 95% CI [−0.18, 0.15], $p = .87$), Updating × Strategy Used ($\beta = -0.08$, 95% CI [−0.25, 0.09], $p = .34$), and Grade × Updating × Strategy Used ($\beta = -0.05$, 95% CI [−0.22, 0.11], $p = .52$).

Discussion

In this study, we examined the contributions of children's working memory updating to their arithmetic performance and behavior with a strategy perspective. Large samples of third and fourth graders were asked to estimate sums of two-digit addition problems by rounding both operands either down or up to the nearest decades. Children's strategy selection and strategy execution on problems of varying difficulties were examined with a developmental perspective. Within recent decades, various works documented age improvements in children's strategy selection and execution (for overviews, see Lemaire, 2017; Siegler, 2007). In the current study, we replicated findings on item-related effects as well as age-related improvements in children's better strategy selection and strategy execution. Like in previous studies, we found that age-related changes in better strategy selection were modulated by the type of problems children solved.

Most important and unique, we examined how children's strategy use was influenced by working memory updating processes. Previous research found that executive functions influence children's strategies as well as age-related changes in strategic behaviors (Ai et al., 2017; Barrouillet & Lépine, 2005; Imbo, Duverne, & Lemaire, 2007; Imbo & Vandierendonck, 2007; Lemaire & Lecacheur, 2011). However, no previous data documented how efficiency of working memory updating processes influences children's arithmetic strategy use. The current data showed that efficient updating contributes to children's better strategy selection and strategy execution and that the relations between children's arithmetic performance and updating change with children's age and problem characteristics. We next discuss the implications of these findings.

Role of working memory updating processes in children's arithmetic strategy use

The current results show that individual differences in children's working memory updating processes contribute to three aspects of arithmetic strategy use: strategy flexibility, better strategy selection, and strategy execution.

Regarding children's strategy flexibility, large samples of third and fourth graders enabled us to identify two qualitatively distinct subgroups. In one group, showing strategy variability, children used the two available strategies in a flexible manner. In the other, smaller group, children used a dominant strategy on all or nearly all problems. Results revealed that children with more efficient updating were slightly more likely to use the two available strategies. This might be best explained by assuming that more EC resources facilitated the execution of processes crucial for trial-by-trial strategy selection (i.e., reactivating both strategies in working memory, analyzing problem features to determine which strategy is the best, and choosing the better strategy before executing it). In a study by Lemaire, Luwel, and Brun (2017), fifth and seventh graders were asked to estimate sums in conditions where only one strategy was available or two strategies were available. On each problem, children were told which strategy to execute. Results showed that children found estimates more quickly when they were using a single strategy rather than two strategies. The authors explained these findings in terms of lower demands placed on working memory and EC processes in the one-strategy condition.

In the current study, children with more efficient updating were more likely to select the better strategy on a problem-by-problem basis than children with less efficient updating. This relation was qualified by these children (a) using the rounding-up strategy more often and (b) being specifically adaptive on homogeneous problems. Moreover, we found that children with more efficient updating were specifically faster on (a) larger addition problems and (b) homogeneous problems.

We can only speculate about which mechanisms underlie the contribution of updating to these specific effects. We argue that these findings can be explained by associative processes. As commonly proposed in models of strategies (Payne, Bettman, & Johnson, 1993; Rieskamp & Otto, 2006; Shrager & Siegler, 1998), the mechanisms underlying strategy selection are based on the activation of costs and benefits of available strategies when being presented with a problem. In the current study, the two available strategies were the rounding-down and rounding-up strategies. Regarding cognitive costs, the rounding-up strategy requires more complex processes (i.e., rounding up the first operand, holding the rounded decade in working memory, executing the same procedure for the second operand, and adding the two updated decades) than the rounding-down strategy (i.e., solely adding the decade units). For children with more efficient updating, the more complex procedures of the rounding-up strategy entail costs that are smaller in relation to their available updating resources, whereas for children with less efficient updating, those relative costs are higher. Hence, in the former group the trade-off between relative costs and benefits is resolved in favor of the more complex procedures more often, leading these children to use the rounding-up strategy more adaptively with the benefit of a more accurate estimate.

In addition, the current study showed that main and interaction effects involving problem type were even larger than the corresponding effects involving rounding type; homogeneous problems were more often solved with the better strategy and were solved more quickly than heterogeneous problems, and this problem type effect interacted with age and updating. When being presented with a problem, it highly depends on children's past learning experience whether costs and benefits of strategies are associated with the problem characteristics (Rieskamp & Otto, 2006; Shrager & Siegler, 1984). Children need to learn which strategies are available. Through experience, they learn to associate which strategy is most adaptive for certain problem characteristics, thereby being associated with larger benefits (e.g., that rounding up two-unit digits with unit digits larger than 5 would lead to a more accurate estimate than rounding down), and they experience which rounding strategy demands more cognitive effort, thereby being associated with larger costs. It is assumed that associations between strategies and certain problem features are strengthened as a result of learning experience and repeated use of algorithms (Shrager & Siegler, 1984), which explains the effect of children being more likely to select the better strategy and showing smaller estimation latencies on homogeneous problems than on heterogeneous problems; only for homogeneous problems could children form associations between problem features and the available strategies (i.e., rounding-down and rounding-up strategies), whereas heterogeneous problems would be associated with the mixed-rounding strategy, a strategy that was not allowed in the current study. The finding that children with more efficient working memory updating were specifically more accurate and faster on homogeneous problems is in line with the assumption that the "WM [working memory] system ... utilizes various control processes that are needed to maintain information in WM and to build strong and durable memories in LTM [long-term memory]" (Unsworth, 2019, p. 81). Theoretical models (Barrouillet, Bernardin, & Camos, 2004; Gavens & Barrouillet, 2004) and empirical works (Barrouillet & Lépine, 2005; Unsworth, Brewer, & Spillers, 2013) on the relation between working memory and long-term memory support the notion that working memory processes contribute to associations in long-term memory being established and to existing associations being accessed. Therefore, it can be assumed that children with more efficient working memory have had an advantage in forming associations prior to the estimation strategy experiment and that during the experiment they had an advantage in retrieving those established associations for familiar homogeneous problems. This enabled them to access the associations for the better strategy more often and faster.

Importantly, contributions of updating processes to children's problem type effect seem to differ between grades. In fourth grade, homogeneous problems were on average solved more than 2 s faster than heterogeneous problems, and this time advantage was independent of children's updating efficiency. In contrast, in third grade, the solution time advantage of homogeneous problems depended on children's updating efficiency. For third graders with efficient updating, a large estimation latency advantage of homogeneous problems over heterogeneous problems was found (like in fourth graders). In contrast, for third graders with less efficient updating, estimation times for homogeneous problems were nearly as slow as those for heterogeneous problems. A plausible explanation lies in different levels of associative learning for third versus fourth graders due to differences in formal training and education (Shrager & Siegler, 1984); all fourth graders, even those with less efficient updating, could already establish associations between familiar homogeneous problems and beneficial strategies minimizing the difference between estimates and exact sums because they had practiced rounding and estimation for more than a year. In contrast, for third graders, estimating approximate sums is a relatively novel task. For those with more efficient updating, the limited experience seems to have been sufficient to form associations between problem features and rounding strategies so that they could solve homogeneous problems faster than heterogeneous problems with the unfamiliar requirement of rounding both operands either down or up.

Note that opposite results regarding the interaction between children's updating processes and problem type could also have been expected. Indeed, updating could be expected to be more relevant for harder (i.e., heterogeneous) problems than for simpler (i.e., homogeneous) problems. Recall that children were not allowed to use the mixed-rounding strategy that would be the dominantly associated strategy for heterogeneous problems. As mentioned above, EC processes are likely involved in two different stages during strategy use: (a) in the consolidation of long-term memory associations during prior learning processes and (b) during the processing of the current task. As revealed by the main effects, efficient updating processes facilitate the processing of current problems in general. Furthermore, updating especially facilitates working on homogeneous problems, indicating that updating influenced the past consolidation of long-term associations and facilitates the current retrieval of these associations. Because on heterogeneous trials the dominantly associated strategy must be inhibited to select between the only available rounding strategies here (rounding down and rounding up), inhibition of the mixed-rounding strategies might be most relevant on these heterogeneous problems.

Finally, results revealed that children with more efficient updating slowed down less on larger addition problems than children with less efficient updating. Existing theories on the development of underlying procedures for (simple) addition are somehow inconsistent. It has long been assumed that age-related decreases in solution times for additions result from children using retrieval more often (Ashcraft, 1982; Siegler, 1987). This has lately been challenged by the assumption that addition always stems from unconscious counting procedures (Thevenot, Barrouillet, Castel, & Uittenhove, 2016). Even though there still remain open questions regarding underlying procedures for addition, our results indicate that updating processes facilitate children's addition specifically for larger problems—whether via the more frequent use of retrieval or more efficient counting procedures.

In sum, the current study is the first to reveal that one EC process that had not been investigated yet, namely updating, plays an important role in children's arithmetic strategies and performance. The findings on the contribution of children's updating processes, especially on complex problems (i.e., rounding up and larger problems), support the notion of working memory to facilitate mental calculations that involve multiple steps (Hitch, 1978). Updating processes, therefore, should play an even larger role in more complex arithmetic that is not limited to two-digit addition problems estimation.

Implications

The findings of the current study have important theoretical implications for strategy choice models. Existing models (Payne et al., 1993; Rieskamp & Otto, 2006; Shrager & Siegler, 1998) share common assumptions such as the activation of associated costs and benefits during strategy selection. However, the models do not assume that strategy selection requires EC processes. Various works have already suggested that cognitive components, such as inhibition and shifting, should be added (Lemaire & Lecacheur, 2011; Luwel, Schillemans, Onghena, & Verschaffel, 2009). The current study strengthens this case by revealing that another executive function, namely working memory updating, contributes to children's better strategy selection. Although the underlying mechanisms of the link between updating processes and strategy selection and execution need further research, it can be speculated that advantages in the consolidation and retrieval of associated costs and benefits of available strategies play an important role. We argue that updating could be included in existing models as influencing the relative importance or weights of associated costs and benefits to the person, similar to weighted additive heuristics (see also Payne et al., 1993). That is, associated costs due to the complexity of strategy operations can be compensated by efficient cognitive processes such as updating.

Together with previous studies showing effects of other EC processes on children's strategy use (e.g., Barrouillet & Lépine, 2005; Lemaire & Lecacheur, 2011), the current findings raise several issues that need to be addressed in future research. One major limitation of the current and previous works is that the data cannot rule out the possibility that other executive functions or general cognitive abilities are involved in strategies and strategic changes. Indeed, several studies with participants of different age groups could consistently find that the performance in each executive function task can be decomposed into a contribution of the so-called *common executive function* and a unique contribution of the particular executive function (see also Miyake & Friedman, 2012). That is, when assessing updating, the score contains variance best explained by a general factor that is common to all executive functions and variance that is unique to updating. Miyake and Friedman (2012) defined the common executive function as the "ability to actively maintain task goals and goal-related information and use this information to effectively bias lower-level processing" (p. 11). This ability is essential to enable participants to engage in other specific processes. Hence, it is fair to expect that some variance in our updating score is shared with a common executive function factor underlying interindividual differences. In the current data, we cannot differentiate between these sources of variance or test the assumption. However, we expect specific executive components to be more relevantly involved in different subprocesses of arithmetic problem solving than others. Because updating processes contain the active manipulation of information in working memory as well as the access to information from long-term memory, updating-specific abilities should be specifically required when retrieving strategy–problem associations from long-term memory and mentally rounding and adding up manipulated operands within a computational estimation task. In turn, the common factor should be the key to maintain the available strategies and inhibit invalid strategies such as mixed rounding for heterogeneous problems.

Going on from this, it would be important to empirically assess the relative contributions of distinct EC processes (shifting, inhibition, and updating) to strategy selection and performance as well as how these contributions change with age. Furthermore, it would be important to investigate whether different EC processes account for different experimental effects such as efficient shifting processes specifically accounting for smaller strategy switch costs or repetition benefits during strategy execution and updating processes specifically accounting for differences in strategy selection due to associative procedures. From a theoretical perspective, such comparative research could answer whether strategy models should include one or several parameters for EC processes to account for individual differences in strategy selection and execution.

Acknowledgment

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sector.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2019.04.003.

References

Ai, J., Yang, J., Zhang, T., Si, J., & Liu, Y. (2017). The effect of central executive load on fourth and sixth graders' use of arithmetic strategies. *Psychologica Belgica, 57*, 154–172.
Ashcraft, M. H. (1982). The development of mental arithmetic: A chronometric approach. *Developmental Review, 2*, 213–236.
Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General, 133*, 83–100.
Barrouillet, P., & Lépine, R. (2005). Working memory and children's use of retrieval to solve addition problems. *Journal of Experimental Child Psychology, 91*, 183–204.
Cohen Kadosh, R., & Dowker, A. (Eds.). (2015). *The Oxford handbook of numerical cognition*. Oxford, UK: Oxford University Press.
DeStefano, D., & LeFevre, J.-A. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology, 16*, 353–386.
Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*, 135–168.
Dirk, J., & Schmiedek, F. (2016). Fluctuations in elementary school children's working memory performance in the school context. *Journal of Educational Psychology, 108*, 722–739.
Friso-van den Bos, I., van der Ven, S. H. G., Kroesbergen, E. H., & van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review, 10*, 29–44.
Gavens, N., & Barrouillet, P. (2004). Delays of retention, processing efficiency, and attentional resources in working memory span development. *Journal of Memory and Language, 51*, 644–657.
Gilmore, C., Göbel, S. M., & Inglis, M. (Eds.). (2018). *An introduction to mathematical cognition*. London: Routledge.
Hitch, G. J. (1978). The role of short-term working memory in mental arithmetic. *Cognitive Psychology, 10*, 302–323.
Imbo, I., Duverne, S., & Lemaire, P. (2007). Working memory, strategy execution, and strategy selection in mental arithmetic. *Quarterly Journal of Experimental Psychology, 60*, 1246–1264.
Imbo, I., & Vandierendonck, A. (2007). The development of strategy use in elementary school children: Working memory and individual differences. *Journal of Experimental Child Psychology, 96*, 284–309.
Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434–446.
Lechuga, M. T., Pelegrina, S., Pelaez, J. L., Martin-Puga, M. E., & Justicia, M. J. (2016). Working memory updating as a predictor of academic attainment. *Educational Psychology, 36*, 675–690.
Lee, K., Ng, S. F., Bull, R., Pe, M. L., & Ho, R. H. M. (2011). Are patterns important? An investigation of the relationships between proficiencies in patterns, computation, executive functioning, and algebraic word problems. *Journal of Educational Psychology, 103*, 269–281.

LeFevre, J.-A., Greenham, S. L., & Waheed, N. (1993). The development of procedural and conceptual knowledge in computational estimation. *Cognition and Instruction, 11*, 95–132.

LeFevre, J.-A., Skwarchuk, S.-L., Smith-Chant, B. L., Fast, L., Kamawar, D., & Bisanz, J. (2009). Home numeracy experiences and children's math performance in the early school years. *Canadian Journal of Behavioural Science, 41*, 55–66.

Lehto, J. (1996). Are executive function tests dependent on working memory capacity? *Quarterly Journal of Experimental Psychology A, 49*, 29–50.

Lemaire, P. (Ed.). (2017). *Cognitive development from a strategy perspective: A Festschrift for Robert Siegler*. London: Routledge.

Lemaire, P., & Brun, F. (2014). Effects of strategy sequences and response–stimulus intervals on children's strategy selection and strategy execution: A study in computational estimation. *Psychological Research, 78*, 506–519.

Lemaire, P., & Brun, F. (2016). Age-related differences in children's strategy repetition: A study in arithmetic. *Journal of Experimental Child Psychology, 150*, 227–240.

Lemaire, P., & Lecacheur, M. (2011). Age-related changes in children's executive functions and strategy selection: A study in computational estimation. *Cognitive Development, 26*, 282–294.

Lemaire, P., Lecacheur, M., & Farioli, F. (2000). Children's strategy use in computational estimation. *Canadian Journal of Experimental Psychology, 54*, 141–148.

Lemaire, P., Luwel, K., & Brun, F. (2017). Does the number of available strategies change how children perform cognitive tasks? Insights from arithmetic. *Journal of Educational and Developmental Psychology, 7*(2), 43–50.

Lemaire, P., & Reder, L. (1999). What affects strategy selection in arithmetic? The example of parity and five effects on product verification. *Memory & Cognition, 27*, 364–382.

Lovett, M. C., & Anderson, J. R. (1996). History of success and current context in problem solving. *Cognitive Psychology, 31*, 168–217.

Luwel, K., Schillemans, V., Onghena, P., & Verschaffel, L. (2009). Does switching between strategies within the same task involve a cost? *British Journal of Psychology, 100*, 753–771.

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*, 8–14.

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

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology, 81*, 111–121.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge, UK: Cambridge University Press.

Rabbitt, P. (Ed.). (1997). *Methodology of frontal and executive function*. Hove, UK: Psychology Press.

Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General, 135*, 207–236.

Shrager, J., & Siegler, R. S. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 229–287). Hillsdale: Lawrence Erlbaum Associates.

Shrager, J., & Siegler, R. S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science, 9*, 405–410.

Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General, 116*, 250–264.

Siegler, R. S. (2007). Cognitive variability. *Developmental Science, 10*, 104–109.

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology, 59*, 745–759.

Steele, F. (2009). *Module 7: Multilevel models for binary responses: Concepts* Retrieved from. Bristol, UK: Centre for Multilevel Modelling, University of Bristol http://www.bristol.ac.uk/media-library/sites/cmm/migrated/documents/7-concepts-sample.pdf.

Thevenot, C., Barrouillet, P., Castel, C., & Uittenhove, K. (2016). Ten-year-old children strategies in mental addition: A counting model account. *Cognition, 146*, 48–57.

Unsworth, N. (2019). Individual differences in long-term memory. *Psychological Bulletin, 145*, 79–139.

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2013). Working memory capacity and retrieval from long-term memory: The role of controlled search. *Memory & Cognition, 41*, 242–254.

van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., & Leseman, P. P. M. (2012). The development of executive functions and early mathematics: A dynamic relationship. *British Journal of Educational Psychology, 82*, 100–119.

6.2 Study 2: Mixed-rounding Strategy Use

Poloczek, S., **Hammerstein, S.**, & Büttner, G. (2020). Children's mixed-rounding strategy use in computational estimation. *Manuscript submitted for publication*.

Children's Mixed-rounding Strategy Use in Computational Estimation

Sebastian Poloczek[1,2*], Svenja Hammerstein[1,2], & Gerhard Büttner[1,2]

[1]Goethe Universität Frankfurt am Main, Germany

[2]Centre for Individual Development and Adaptive Education of Children at Risk (IDeA), Frankfurt am Main, Germany

*Corresponding author at: Goethe Universität Frankfurt am Main, Theodor-W.-Adorno-Platz 6, 60629 Frankfurt am Main; Tel.: +49 69 798 35380; E-mail address: poloczek@paed.psych.uni-frankfurt.de (P. Poloczek)

Abstract

Being able to efficiently perform computational estimations is an important skill. Furthermore, computational estimation is used to study strategy development and the role of executive functions therein. However, little is known about children's mixed-rounding strategy use. Additionally, the design of many studies examining the involvement of executive functions in computational estimation restricted participant's strategy choice to rounding both operands down or rounding both up, even though mixed-unit problems were present. Preventing children from using the potentially preferred mixed-rounding strategy and requiring them to consider unit sums in the strategy decision could artificially increase demands on executive functions. Therefore, we presented fourth graders with addition estimation problems, in which the size of both unit digits was systematically varied, without preventing mixed-rounding and additionally assessed the involvement of executive functions. Most children adjusted their strategy choice to the unit digits, therefore, preferring mixed-rounding on mixed-unit problems. Additionally, the sum of units had only little influence on children's strategy choice, even though considering this sum would be required when mixed-rounding is excluded. These results suggest that the design preventing participants to use mixed-rounding might increase task demands. This conclusion was further supported by not finding strategy switching costs (explained by demands on executive functions) and by failing to find effects of working memory or shifting capabilities on how well participants adjusted their strategy choice to the problem. Consequently, the role of executive functions in computational estimation might have been overestimated by previous research and further research should be done with more realistic designs.

Keywords: arithmetic; computational estimation; strategies; children; working memory; shifting

Introduction

Being able to correctly and efficiently perform computational estimations is an important skill in everyday life and in learning mathematics. The ability allows children to make approximate calculations without the need for a calculator or paper and pencil, both in real-world situations and to check the reasonableness of complex calculations found through other means. Additionally, it may help them to “develop a better understanding of place value, mathematical operations, and general number sense” (Star, Rittle-Johnson, Lynch, & Perova, 2009, p. 569). Computational estimation tasks can be broken down into (at least) two subtasks: approximation and mental computation (LeFevre, Greenham, & Waheed, 1993). As an example, to estimate the sum of 54 and 87, first, students have to reformulate the task, e.g. by rounding the first number down and rounding the second number up to the nearest whole 10; second, they have to perform a mental calculation on the approximate numbers to derive an estimate of e.g. 140. The approximation subtask offers some degrees of freedom so that different approximations can be made and therefore, different computational estimation strategies are available.

In good computational estimators, three key processes or broad categories of strategies were identified for the approximation step: reformulation, translation, and compensation (Reys, Rybolt, Bestgen & Wyatt, 1982). *Reformulation* refers to simplifying the numerical data to produce a mentally more manageable problem, like rounding or truncating the numbers in the estimation problem. *Translation* refers to altering the mathematical structure to create a mentally more manageable problem, e.g. translating 83 + 74 + 96 + 82 into about 4 x 80 is roughly 320. *Compensation* refers to adjustments made during intermediate steps or after the mental calculation. Compensation reduces the discrepancy between the estimate and the exact calculation, like increasing the estimate in the previous example from 320 to 330 as 3 out of 4 operands were larger than 80.

A variety of strategies for computational estimation have been documented. Most studies agree that children and adults typically use multiple strategies and that most trials are solved with some sort of rounding strategy, even though exact strategy classifications vary between studies (e.g. LeFevre et al., 1993; Lemaire, Lecacheur, & Farioli, 2000; Star & Rittle-Johnson, 2009; Star et al., 2009; Xu, Wells, LeFevre, & Imbo, 2014). Truncation (i.e., replacing the unit digits with 0) was also present but after the introduction of rounding rules less frequent than adaptive rounding (Hammerstein, Poloczek, Lösche, Lemaire & Büttner, 2019; Lemaire et al., 2000; Star & Rittle-Johnson, 2009, Xu et al., 2014). Compensation approaches were only rarely used but become more frequent with increasing age (LeFevre et al., 1993; Lemaire et al., 2000; Xu et al., 2014).

Children not just use multiple strategies, but they typically adapt their strategy choice to features of the estimation problem. Rounding strategies can be divided into three types: rounding-down (i.e., all operands are rounded down), mixed-rounding (i.e., some operands are rounded down and the others are rounded up), and rounding-up (i.e., all operands are rounded up). They favour the rounding-down strategy for problems, for which rounding-down provides a close estimate, and they prefer rounding-up for those problems, for which the rounding-up result is closer to the exact calculation (e.g. Lemaire & Lecacheur, 2002). Another example of adaptivity is that children rely more on their sense of magnitude and less on approximate calculation involving rounding if estimates and reference numbers were far apart vs. close (Ganor-Stern, 2015, 2016). The quality and efficiency of adaptive strategy choices were linked to children's arithmetic skills (Dowker, 1997; LeFevre et al. 1993; Seethaler & Fuchs, 2006). Interestingly, they seem to be linked to executive functions like inhibition and shifting (Lemaire & Lecacheur, 2011) and working memory updating (Hammerstein et al., 2019; Seethaler & Fuchs, 2006). Executive functions are regulatory mechanisms of the mind with three correlated, but separable functions: *inhibition* as the capability to deliberately inhibit dominant responses; *shifting* as the capability to shift between mental sets or multiple tasks and *working memory updating* as the capability to store, manipulate, monitor and update information (compare Miyake et al., 2000).

So far only extremely limited, systematic knowledge exists on the adaptive use of the mixed-rounding strategy in children for two reasons. Firstly, in several studies (typically using tasks with two operands) participants' choice of strategies was restricted to rounding either both operands down or rounding both operands up (e.g., Ai et al., 2017; Hammerstein et al., 2019; Lemaire & Brun, 2016; Lemaire & Lecacheur, 2011; Lemaire et al., 2017). This restriction was in place even for problems in which mixed-rounding would yield the closest estimate (e.g., being asked to estimate 32 + 57 by calculating 30 + 50 or 40 + 60, but not mixed-rounding with 30 + 60). Secondly, studies not restricting strategy choice typically did not distinguish between rounding-down, mixed-rounding, and rounding-up (Dowker, 1997; Ganor-Stern, 2016;

Lemaire, et al., 2000; Seethaler & Fuchs, 2006; Star & Rittle-Johnson, 2009; Star et al., 2009) so that little is known about the use of mixed-rounding. A rare exception were experiments with undergraduates (Xu et al., 2014) showing that mixed-rounding was roughly as common as either rounding-up or rounding-down. However, it is not clear, whether the results generalize to school-aged children.

Studying mixed-rounding use in school-aged children in detail is important for at least two reasons. *Firstly,* when rounding is taught at school mixed-rounding is an integral part of what children learn. To better understand children's adaptive use of the whole range of different rounding strategies, it is necessary to investigate mixed-rounding use in addition to rounding-down and rounding-up. *Secondly,* the ecological validity and generalizability of some results on computational estimation obtained with the restricted design with mixed-rounding excluded is questionable. This restriction was deliberate "to increase difficulty of strategy choices and, thereby, to avoid ceiling effects" (e.g. Lemaire & Brun, 2014, p. 510) as this would maximise the chances to detect sequential effects like switching costs and related age-group differences. While this was a reasonable design choice and led to significant findings on the afore mentioned aspects of rounding, we argue that the restriction potentially limits the ecological validity and generalizability especially of results on the relationship between estimation strategy use and executive functions. The reason for this consideration is that additional cognitive demands are present in estimation tasks with a restricted compared to the unrestricted choice of rounding strategies. In a restricted design, problems for which mixed-rounding would lead to the closest estimate have to be solved with an unintuitive approach if participants prefer mixed-rounding for these problems. They have to inhibit their first choice (which would be the optimal problem-based strategy; cf. Xu et al., 2014) and figure out whether rounding-down or rounding-up is the second-best strategy. Most importantly, to do so, looking at both unit digits individually is not sufficient, but the sum of unit digits has to be considered. That is, for mixed-unit problems with unit sums below 10 like 32 + 57 rounding-down would yield the second closest estimate (since the estimate of 80 is closer to the exact sum of 89 than the estimate of 100) and for problems with unit sums above 10 like 34 + 57 rounding-up would yield the second closest estimate (since this time the estimate of 100 is closer to the exact sum of 91 than the estimate of 80). Figuring out unit sums requires additional mental processes with working memory involvement. Based on these considerations we argue that the artificial exclusion of mixed rounding could lead to additional demands on executive functions that are not genuine to a not restricted, more naturalistic setting of computational estimation. Against this background, the question arises as to the generalizability of previous findings on the relationship between estimation strategy use and executive functions particularly as nearly all studies examining the involvement of executive functions in computational estimation applied the artificially restricted design with only two strategies (for inhibition and shifting see Hodzik & Lemaire, 2011; Lemaire & Lecacheur, 2011; for switching costs - these are the costs associated with changing strategy over consecutive trials that are explained with demands on the executive function shifting - see Lemaire & Lecacheur, 2010; for working memory see Ai, Yang, Zhang, Si, & Liu, 2017; Hammerstein et al., 2019; only exception for working memory involvement in a non-restricted design see Seethaler & Fuchs, 2006).

The Present Study

The present study focussed on two different aspects of mixed-rounding. In the first part, we investigated the use of mixed-rounding in primary school children in estimation problems when rounding-down, rounding-up and mixed-rounding as well was allowed. Fourth graders completed computational estimation problems for two-digit additions by rounding-down, rounding-up or mixed-rounding. Unit sizes were systematically varied, resulting in equal numbers of otherwise comparable small-unit problems, large-unit problems and mixed-unit problems. Data screening showed that children could approach estimation trials in an inflexible manner sticking to the same strategy or a flexible manner using different rounding strategies for different estimation problems. We focussed on the children who tended to use more than one rounding strategy. First, we were interested in the pattern of their strategy use and analysed whether these children systematically adapted their choice of a rounding strategy to the main categories of small-unit, mixed-unit and large-unit problems (RQ 1a). It was of special interest whether children preferred mixed-rounding for mixed-unit problems.

Second, we were interested in a more detailed analysis. A crucial feature of the design was that within the main problem categories sums of both unit digits were systematically varied. We analysed whether the sum of units did influence strategy choice once the effect of the main categories was controlled for (RQ 1b). The rationale for this research question was that in designs with a restriction to rounding-down and rounding-up strategies (as used in many previous studies), mixed-unit problems with sums of units less than 10 are best solved by rounding-down whereas in problems with sums of units more than 10 rounding-up leads to the closest estimate. That means to find the best solution participants are forced to consider the sums. In contrast, in designs with no restriction in the use of rounding strategies it is not necessary to compute the unit sums for mixed-unit problems. But if children strive to give close estimates, they could consider unit sums when deciding for a strategy on small-unit or large-unit problems. This is because the closest estimate for small-unit problems with unit sums adding up to more than 5 (e.g. 53 + 74) or for large-unit problems adding up to less than 15 is given by using mixed-rounding or compensation. As stated above considering sum of units requires additional mental processes which may lead to a noticeably different cognitive load when solving mixed-unit problems in restricted versus non-restricted designs. Especially if children adaptively prefer mixed-rounding for mixed-unit problems, but typically do not consider the sum of units in their choice, the ecological validity of the design preventing mixed-rounding is questioned.

Thereafter, we examined in a second part of the study whether effects on best strategy choice and estimation latencies (an approximation of the efficiency of strategy use) reported in the literature do generalize to a design in which mixed-rounding is allowed. We investigated the following three research questions:

Do working memory and shifting capabilities predict differences in how well children adapt their strategy choices to the estimation problems and in how fast they come up with good estimates when differences in children's arithmetic skills are controlled (RQ 2a)?

Can switching costs be detected as these costs are explained as stemming from the demand on executive functions when switching between tasks (RQ 2b)?

Do problem effects not directly linked to executive functions generalize? Do children use the best strategy more often on small-unit problems compared to mixed-unit or large-unit problems and is rounding-down faster than mixed-rounding and rounding-up? Does problem size affect estimation latencies and the likelihood of using the best strategy (RQ 2c)?

We expected that effects not directly linked to executive functions should be replicated. But the focus of the second part was whether effects of executive functions on strategy choice and estimation latencies do generalize.

Method

Participants

Eighty-eight children were tested (46 males; age in months: *M* = 122.0, *SD* = 5.4, range = 111-138). All participants were attending the second half of 4th grade. They were recruited from eleven classes in seven primary schools in urban and suburban areas in the state of Hesse (Germany). The study was approved by the local ethics committee. Parents provided their written informed consent, and children gave their verbal consent.

Tasks

Computational estimation experiment.

Pool of estimation problems. Computational estimation problems were drawn from three main categories: 24 small-unit problems with unit digits of both operands smaller than 5; 24 mixed-unit problems with one of the unit digits smaller and the other larger than 5; and 24 large-unit problems with unit digits of both operands larger than 5.

For two reasons, each of those three categories were further subdivided into three subcategories according to the sums of the unit digits. First, unit sums are decisive for which rounding strategy leads to the closest estimate. Second, in previous research, unit sums determined for mixed-unit problems whether a problem was classified as heterogeneous small problem vs. heterogeneous large problem with rounding-down vs. rounding-up as (second) best strategy (see Table 1). It is essential to answer the question whether children consider the sum of the unit digits in their strategy choice once the effect of the three main categories is taken into account.

- The small-unit problems were subdivided into *small1* for rounding-down problems with unit digits adding up to 3 or 4, thus, rounding-down being unambiguously the best strategy yielding the estimate closest to the exact sum; *small2* for problems with both unit digits smaller than 5 suggesting rounding-down, but adding up to 5, thus, rounding-down and mixed-rounding resulting in equally close estimates; *small3* for problems with both unit digits smaller than 5 suggesting rounding-down, but adding up to 6 or 7, thus, mixed-rounding yields the closest estimate.

- The mixed-unit problems were subdivided into *mixed1* for problems with unit sums between 7 and 9 (note that these problems were classified as heterogeneous small-unit problems in previous studies in which mixed-rounding was forbidden); *mixed2* for problems with units adding up to 10; *mixed3* for mixed-unit problems with unit sums between 11 and 13 (note that these problems were previously classified as heterogeneous large-unit problems in studies in which mixed-rounding was forbidden).
- The subcategories for large-unit problems were as following: *large1* with both unit digits above 5 but unit sums of 13 and 14, thus with mixed-rounding providing the closest estimate; *large2* with units adding up to 15, therefore mixed-rounding and rounding-up resulting in equally close estimates; and finally *large3* with unit sums of 16 and 17 and rounding-up unambiguously the best strategy.

Table 1

Illustration how estimation problems were organised in main categories and subcategories and how these categories relate to a classification in previous studies

Main category defined by size of unit digits	**Subcategory** Name	defined by sum of units	Examples	**Closest estimate** by	**Classification** in previous research
small-unit both smaller than 5	small1	3 or 4	_1 + _2, _1 + _3	rounding-down	homogeneous small-unit problems
	small2	5	_1 + _4, _2 + _3	mixed-rounding	
	small3	6 or 7	_2 + _4, _3 + _4		
mixed-unit one smaller than 5, the other larger than 5	mixed1	7, 8 or 9	_1 + _6, _2 + _7		heterogeneous small
	mixed2	10	_3 + _7, _2 + _8		not included
	mixed3	11, 12 or 13	_3 + _8, _4 + _9		heterogeneous large
large-unit both larger than 5	large1	13 or 14	_6 + _7, _6 + _8		homogeneous large-unit problems
	large2	15	_6 + _9, _7 + _8	rounding-up	
	large3	16 or 17	_7 + _9, _8 + _9		

To avoid systematic errors in the composition of the operands for half of the problems in each subcategory the unit digit of the first operand was larger than the unit digit of the second operand. The pool of unit digit pairs (e.g., _1 + _4, _7 + _2) was combined with the pool of expected additions after rounding (e.g., 30 + 60, 80 + 50) with the constraints that for each subcategory 50% of additions were without carry (estimates of 50-100) and 50% with carry (estimates of 110-170) and that in 50% the first operand was larger than the second one. Additionally, the sums of the exact calculations and the estimates given the best rounding strategy were matched as closely as possible for the subcategories. Further constraints comparable to previous research (Lemaire & Brun, 2016; Lemaire & Lecacheur, 2011) were applied when constructing the items. Additionally, items were put into a tightly controlled pseudorandom sequence with about 50% main task category repetitions and 50% task category switches (for further details see supplementary material 1.1 and 1.2).

Procedure of the computational estimation task. Children were asked to give an approximate answer for two-digit additions without calculating the exact sum. With the example 28 + 41 different possibilities to get to an estimate were introduced: rounding down both operands to 20 and 40 and giving 60 as answer (rounding-down strategy); rounding up both operands to 30 and 50 with 80 as answer (rounding-up strategy); and rounding one operand up while rounding the other one down to 30 and 40 with 70 as estimate (mixed-rounding strategy). Participants were told that they could decide how they estimate the result but that they should do it in a way that yields estimates close to the exact sum while being fast at the same time. Note, that in contrast to most previous studies, mixed-rounding was included. Children were not instructed to indicate which strategy they were using, but strategy selection was inferred from the responses to the estimation trial (e.g., for the trial 22 + 57, a child was considered to have used the rounding-down strategy after providing 70 as response, the mixed-rounding strategy after providing 80, and the rounding-up strategy after providing 90; for further details, see supplementary material 1.3).

The task was divided into two test sets on different days. For each of the two sets, children were administered practice trials with adaptive feedback. After instruction and practice, no participant displayed any apparent difficulties with the task. Estimation problems were presented in black (Font: Arial, Font size: 150) on an otherwise white screen, until the participant had typed in their answer. No time-limit for responding was imposed. Within each set of the estimation task, children completed two blocks of 18 trials; therefore, in total 72 estimation trials.

Assessment of cognitive capabilities. Working memory updating, shifting and inhibition are all separable, but still correlated executive functions (e.g. Miyake et al., 2000), and all executive function are correlated to performance in mathematics including simple arithmetic skills (see Friso-van den Bos, van der Ven, Kroesbergen, & van Luit, 2013). To disentangle the potential effects of different cognitive capabilities on computational estimation and to address the task-impurity problem, we therefore assessed three executive functions and arithmetic fluency with multiple tasks (for more detailed descriptions of tasks, including reliability calculations, see supplementary material 2.1-2.5).

Working memory. The capability to store and manipulate information in working memory was assessed by three subtests. Participants were administered a *Spatial Updating* task (Dirk & Schmiedek, 2016) in which they were presented with coloured, fictitious creatures at different positions on a grid. Participants were asked to remember and update the positions of two or three creatures that were covertly moving on the grid. In the *Colour Updating* task, participants were visually and verbally presented with sequences of colours, one at a time. When the presentation stopped, children were asked to report the last two or three colours in correct forward order. In the *Picture Span Backward* task, trials consisted of a sequences of nameable line drawing. Participants were asked to reproduce the sequence in reverse order on a response board with 9 pictures.

Shifting. A second component of executive functions, the capability to shift between tasks was assessed by the *Animal Colour* task (Kray, Eber, & Karbach, 2008) and by the *Colour Shape* task (Espy, 1997). Both tasks started with two single task blocks, in which participants had to sort stimuli as fast as possible according to one decision rule (flying or swimming animals, colour of the picture, shape of the figure). During mixed blocks, a cue indicated the rule according to which stimuli had to be sorted; sometimes participants had to switch between the two rules. The differences between the reaction times in the mixed task blocks and the single task blocks were taken as shifting indicators. Thus, higher scores indicate larger switching and mixing cost or lower shifting capabilities.

Excluded executive function tasks. One shifting task and four inhibition tasks were excluded from further analyses due to a lack of reliability. Problems regarding the reliability of inhibition scores as individual difference measures are not specific to this study but are starting to be acknowledged and reported more widely (Hedge, Powell, & Sumner, 2018).

Arithmetic fluency. For arithmetic fluency, we assessed (sub)skills needed to solve computational estimation problems: to decide, whether a number should be rounded up or down, to carry out rounding and to add numbers. The *Speed of Rounding Decisions* task required children to indicate whether two-digit numbers should be rounded down (e.g., 43) or up (e.g., 68). During the *Speed of Cued Rounding* task, two-digit numbers appeared on the screen while arrows indicated in which direction to round. Participants were asked to enter their response as fast as possible on the number pad. Children's *Addition Speed* was operationalized by how fast they correctly solved 24 one- and two-digit addition problems (e.g., 5 + 8, 30 + 60); half of them with and half without carry.

Procedure of the Study

All participants completed a computational estimation task of addition problems divided into two sets, ten executive function tasks, and three arithmetic fluency tasks. On four different days (average time between first and last session: 15.7 days, range: 6-32 days), children participated in four sessions, each lasting approximately 45 minutes. Within each session, arithmetic as well as executive function tasks were mixed, but tasks were administered in the same order for all participants.

Children were tested in groups of up to 5 children. They solved all tasks, programmed with E-Prime, on computers. Instructions were presented verbally over headphones and additional instructions were given if children did not respond or made errors during the practice trials. Two experimenters were present for further questions and instructions.

Children responded by typing their answers. Therefore, at the start of the study, children were familiarized with the number pad. To not influence rounding strategy use, the tasks Speed of Rounding Decision and Speed of Cued Rounding were presented after the two sets of the Computational Estimation tasks.

Data Analyses

Data of the estimation experiment had a cross-classified structure with trials nested in items and in participants. To examine effects of participant characteristics and effects of problem features while taking random effects of participants and random effects of items into account, data were analysed with cross-classified multilevel models or (generalized) linear mixed models ((G)LMM). The advantages of (G)LMMs compared to ANOVAs include that logistic GLMMs avoid known problems that occur when proportion data are analysed with ANOVAs. In addition, in (G)LMMs both categorical and continuous predictors at the participant level (e.g. working memory), at the item level (e.g. main category or problem size), and even at the trial level (e.g. the strategy used on a given trial like mixed-rounding or an estimation strategy switch) can be analysed jointly (see Hofmann & Rovine, 2007; Jaeger, 2008). All logistic GLMMs for categorical dependant variables like strategy choice and best strategy choice and the LMMs for continuous estimation latencies were performed with MLwiN 3.02 (Charlton, Rashbach, Browne, Healy, & Cameron, 2018) with MCMC estimation (Browne, 2017), with 100000 iterations and thinning to 5000 estimates from R with the R2MLwiN package (Zhang, Parker, Charlton, Leckie, & Browne, 2016).

Part 1: Use of Mixed-rounding

Results

Strategic flexibility. Data inspection showed that some children used the same rounding strategy on (almost) all computational estimation problems within the same task block, while other children switched between different rounding approaches. To systematically analyse this, the proportion of trials solved with the most common strategy of this block was computed. The resulting histogram (see supplementary material 1.4) revealed a clearly bimodal distribution with maxima around 40% and at 100% and a minimum around 75%. This suggests that children approached the trials within a block either in a flexible manner, trying to adjust their estimation strategy to the problem characteristics, or in an inflexible way, solving most or all problems with one dominant strategy. A task block was classified as being solved with an inflexible approach, when more than 75% of all (validly classifiable) trials were solved with the same strategy. While 52 children approached all four blocks in a flexible way, 22 children always used an inflexible approach. For the remaining 14 children, some blocks were classified as inflexible and others not. Importantly, for those children the proportions of preferred strategy use ranged across the whole continuum of possible values. Only a few blocks (10 out of 56) had scores of +/-10 percentage points around the 75% cut-off, and no child had one block just below and another one just above the cut-off. This suggests that those 14 children switched between an inflexible approach on some blocks and a flexible approach on others, and that this group was not merely created by unreliability or the choice of the cut-off.

In most blocks with an inflexible approach (102 out of 114), rounding both operands down was the dominant strategy. An inflexible approach with mixed-rounding was very rare (2 out of 114 inflexible blocks).

Flexible strategy selection. Task blocks with an inflexible approach were excluded from the following analyses on flexible strategy selection. Including all children would have been problematic because two subpopulations with distinct approaches and distinct central tendencies of responding were found. The results of the combined analyses would neither be representative of the flexible nor the inflexible approach. Therefore, the sample consisted of 4015 trials from 66 students in blocks with a flexible approach.

These analyses address the first research questions and investigate the effects of main categories (RQ 1a) and subcategories of unit sums (RQ 1b) on children's strategy selection. Data were analysed with cross-classified GLMMs with random intercepts for items and for participants. Strategy selection was analysed with three logistic models: the central one for mixed-rounding and for completeness one for selecting rounding-down and another for rounding-up. Results are displayed in Table 2, however, for a more intuitive understanding of the data see Figure 1 with observed proportions and predicted probabilities of strategy use for different (sub-)categories of estimation problems (for a description of how predicted probabilities including credible intervals were calculated, see supplementary material 3.2). Proportions as well as predicted probabilities of strategy choice show that children selected mixed-rounding on about 84% of mixed-unit trials, while this strategy was only chosen on ~7% of small-unit trials and on ~12% of large-unit trials (compare also very large and significant effect for the mixed-vs.-small-contrast of the mixed-rounding model in Table 2). Likewise, rounding-down was by far the preferred strategy for all small-unit problems (compare also the very large negative effects for mixed-vs.-small and large-vs.-small) and rounding-up the dominant strategy for large-unit problems (compare also the large effect for large-vs.-small).

In addition, it was tested whether subcategories defined by unit sums influences strategy choice. Within small-unit and large-unit problems, *small1* and *large3* were taken as reference, because rounding-down and rounding-up, respectively, are unambiguously the best strategy for items in these subcategories. The only two significant effects out of 12 comparisons were that children were more inclined to choose a mixed strategy on *small2* problems (~9%) and *small3* problems (~7%) than on *small1* problems (~4%). But given that these effects were numerically small and given that all other effects were non-significant despite narrow credible intervals, it can be concluded that strategy choices did not depend substantially on unit sums but on the main categories of small-unit and large-unit. For mixed-unit problems, *mixed2* with unit sums of 10 was taken as reference subcategory. Neither the strategy choices for *mixed1*, in previous studies rounding-down expected, nor the choices for *mixed3*, in previous studies rounding-up expected, differed from *mixed2* problems.

Table 2

Logistic GLMMs for the chosen strategy with main task categories and subcategories according to unit sums as predictors

	Strategy chosen								
	rounding-down			**mixed-rounding**			**rounding-up**		
Fixed Part	**β**	**95% CI**	***p***	**β**	**95% CI**	***p***	**β**	**95% CI**	***p***
Intercept	3.20	[2.65, 3.77]	< .001	-3.34	[-3.91, -2.83]	< .001	-4.49	[-5.27, -3.76]	< .001
Item level predictors									
Mixed (vs. small)	-5.70	[-6.43, -4.99]	< .001	5.02	[4.41, 5.70]	< .001	1.23	[0.41, 2.08]	< .001
Large (vs. small)	-6.68	[-7.49, -5.93]	< .001	1.19	[0.59, 1.85]	< .001	6.45	[5.73, 7.24]	< .001
Small2 (vs. small1)	-0.62	[-1.27, 0.02]	.06	0.90	[0.28, 1.55]	.006	-0.58	[-1.66, 0.47]	.28
Small3 (vs. small1)	-0.60	[-1.26, 0.05]	.07	0.67	[0.03, 1.33]	.05	0.29	[-0.58, 1.20]	.52
Mixed1 (vs. mixed2)	0.19	[-0.48, 0.83]	.57	-0.07	[-0.54, 0.38]	.76	-0.12	[-0.76, 0.52]	.71
Mixed3 (vs. mixed2)	-0.34	[-0.98, 0.31]	.30	0.33	[-0.13, 0.79]	.15	-0.29	[-0.93, 0.36]	.37
Large1 (vs. large3)	0.28	[-0.48, 1.07]	.47	0.21	[-0.27, 0.67]	.38	-0.26	[-0.71, 0.19]	.24
Large2 (vs. large3)	-0.14	[-0.95, 0.66]	.74	0.10	[-0.38, 0.56]	.68	0.04	[-0.42, 0.49]	.84
Random Part	***u***	**95% CI**	**Δ *DIC***	***u***	**95% CI**	**Δ *DIC***	***u***	**95% CI**	**Δ *DIC***
Item intercept	0.15	[0.01, 0.36]	12.3	0.05	[0.001, 0.17]	3.1	0.03	[0.001, 0.13]	-0.9
Participant intercept	1.00	[0.24, 1.55]	192.3	0.19	[0.08, 0.34]	35.0	2.19	[1.40, 3.34]	314.0

Note. 95% CI = 95% credible interval; **Δ** DIC = change in Bayesian Deviance Information Criterion, if random intercept dropped from model.

For further details on how predictors and reference categories were set up, see supplementary material 3.1.

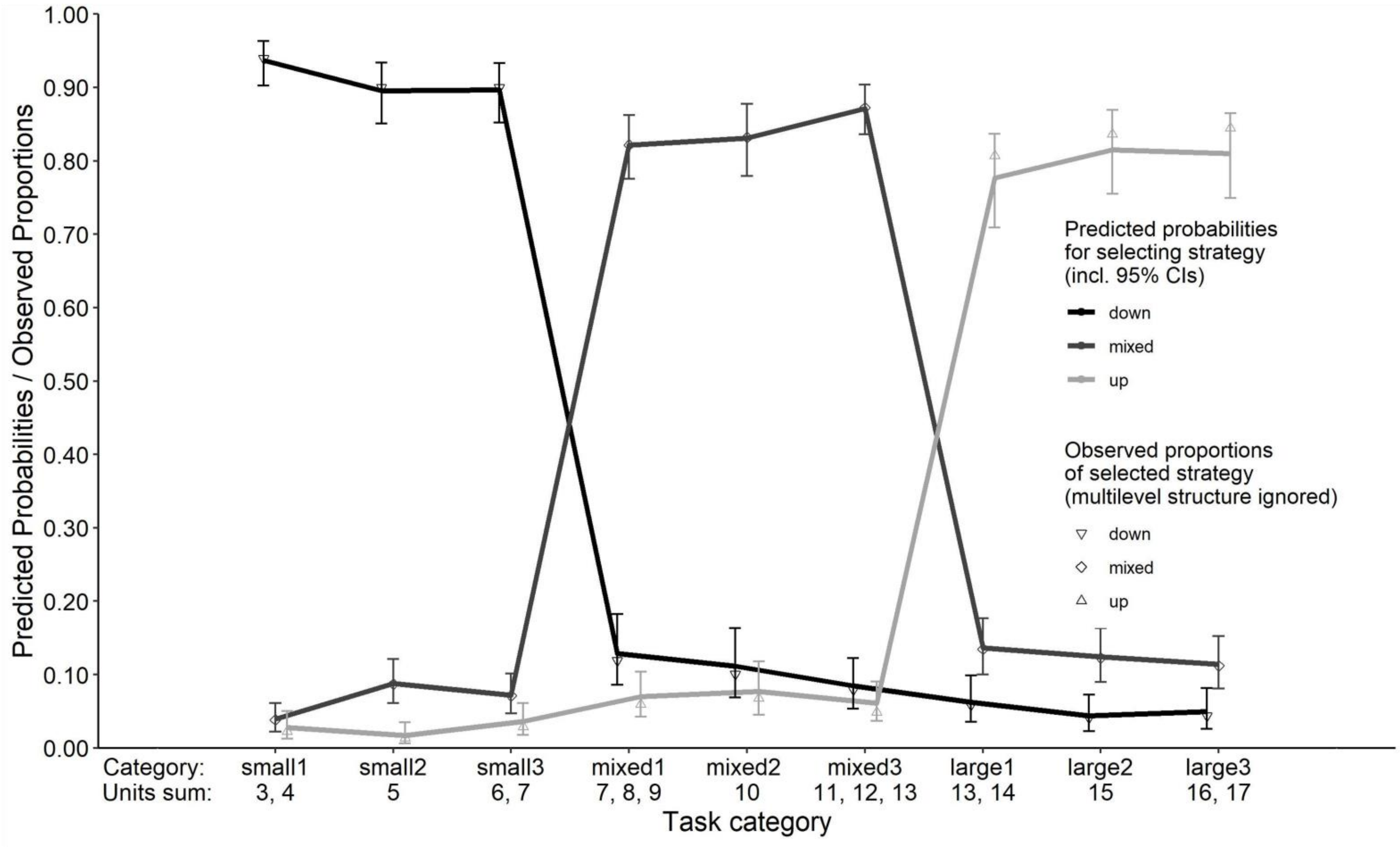


Figure 1. Strategy choice as a function of task category for children with a flexible approach. This plot and the corresponding analyses were based only on task blocks, during which children pursued a flexible approach to strategy selection. If children with an inflexible approach (typically rounding-down irrespective of the size of the units, therefore, truncation) were included, the curves for mixed-rounding and rounding-up would be lower.

Discussion

If children solved all or almost all problems in a test block with the same rounding strategy, all trials of these test blocks were excluded from the main models in order to not mix results of two subpopulations with qualitatively different approaches to the estimation tasks. Many, but not all children adopted a flexible approach with a variety of strategies rather than a single strategy which is in line with previous research (for a review see Siegler & Booth, 2005). Still, about 25% of children were classified as adopting an inflexible approach which was roughly comparable with the numbers reported by Hammerstein et al. (2019). Those opting for an inflexible approach typically did so by choosing rounding-down but applying it irrespective of unit size and therefore using a truncation strategy (Reys et al., 1982). Sticking to mixed-rounding which would result in close estimates for many problems, was very rare.

Fourth graders with a flexible approach used mixed-rounding for about 84% of the mixed-unit problems (RQ1a). Similarly, most small-unit problems were solved by the rounding-down strategy and most large-unit problems by the rounding-up strategy. Therefore, in children with a flexible approach, strategy choice did not vary unsystematically (compare Siegler & Booth; 2005), but children adjusted their strategy to the main category of the problem (i.e. the unit digits of the operands being below or above 5). Importantly, subcategories of problems distinguishing between different unit-sums had no or little impact on strategy choice (RQ 1b3): no differences between different subcategories of mixed- or of large-unit problems were present; for small-unit problems significant, but only numerically small differences were found. Therefore, it can be concluded that fourth graders typically did not consider both unit digits jointly and did not apply prior or post compensation to reduce the rounding distortion of their rounding approach. This is in line with previous research, because compensation was very rarely observed in fifth graders (Lemaire et al., 2000) and because children's insight into the importance of compensation improves with grade level, but the actual use of compensation seems to lag behind recognizing its importance (LeFevre et al., 1993). These findings (strategy choice clearly adjusted to main category like mixed-unit and no substantial effect of unit sums) are important because they question the ecological validity of studies examining computational estimation and excluding mixed-rounding by design. First, mixed-rounding not only provides the closest estimates for mixed-unit problems, which are present in this research tradition, but children would typically opt for mixed-rounding if this strategy is not prevented and has to be inhibited. Second, selecting a second-best strategy requires participants to consider the combined sum of both unit digits, because rounding both operands down is second best with unit sums below 10 and rounding-up with unit sums above 10. However, children of an age often included in these studies do not seem to consider unit sums as compensation was rare. Both findings support the notion that the cognitive demands of computational estimation in the design excluding mixed-rounding while presenting mixed-unit problems are higher than in a design without this restriction. Consequently, it should be examined whether findings of studies with the restricted design generalize. This was the aim of the second part of the present study.

Part 2: Executive Functions in Computational Estimation?

The next section presents effects on best strategy choice and therefore examines which variables predicted how well strategies were adapted to problem features. The subsequent section presents effects on estimation latencies and therefore examines which variables predicted how fast good estimates were given. In both sections, the main interest is on whether children's working memory and shifting capabilities (RQ 2a) and whether switching costs explained by demands on executive functions (RQ 2b) affect strategy choice and estimation latencies in the design with mixed-rounding allowed. Nevertheless, further variables not directly associated with executive functions (RQ 2c) were included for two reasons. Firstly, especially if no significant effects of executive functions are found, it is important to know whether other established effects did replicate or generalize to the present study design to show that the absence of executive function effects was not only due to flaws in the present design.. Secondly, model results are more trustworthy if further important predictors are included instead of omitted.

Results

Effects on selecting the best strategy. Following Xu and colleagues' (2014) rule for balancing simplicity of rounding and computation with proximity of estimates to exact results, we had planned to classify those trials as solved with the best strategy on which the strategy chosen matched the individual rounding rules for each operand. As the findings of the first part suggested that fourth graders selected strategies according to the main problem categories and did not solve small-unit or large-unit problems with unit sums closer to 10 than to 0 or 20 with mixed-rounding and therefore, did not use compensation, this rule for classifying best strategy use appropriately matches children's typical choice.

Children with a flexible approach to strategy selection chose the best strategy on many trials (large positive intercept, compare Table 3 for estimates, CIs and p-values). At the same time, there were clear inter-individual differences in how often children chose the best strategy and therefore how well they adapted their strategy choices (range: 22%-100%; large participant intercept variance). Still, these differences could not be explained by children's executive functions. Neither the effect of working memory (p = .52) nor the effect of shifting capabilities (p = .31) were significant participant level predictors. Equally, no effect of arithmetic fluency (p = .39) was found.

To examine strategy switching cost, estimation problems were arranged in a pseudorandom sequence which ensured that for half of the problems the main task category (small-, mixed-, large-unit) of the preceding problem was repeated and for half the problems it switched. Children were as likely to choose the best strategy as on task-category switch trials as on repetition trials (predicted probability (PP) on switch trials to choose the best strategy: PP_{switch} = 81%, $PP_{repetition}$ = 83%; switch-to-repetition contrast: p = .09). Therefore, no switching cost occurred, and children were as adaptive on trials for which they had to switch strategy if they wanted to choose the best strategy as they were on repetition trials.

Regarding the question, whether other established effects generalize to the current design with mixed-rounding permitted, the results were as following: As expected, in comparison to small-unit trials (PP = 89%), the best strategy was chosen less often on mixed-unit problems (PP = 82%, mixed-to-small contrast: p < .001) and less often on large-unit problems (PP = 79%; large-to-small contrast: p < .001). Additionally, there was the expected effect of problem size (p < .001), because on trials with larger estimates children were somewhat less likely to choose the best strategy compared to trials with smaller estimates ($PP_{estimate=140}$ = 81% vs. $PP_{estimate=70}$ = 86%).

Table 3

Logistic GLMM examining predictors of best strategy use

Fixed Part	**β**	**95% CI**	***p***
Intercept	3.31	[2.76, 3.88]	< .001
Participant level predictors			
Working memory (z-stand.)	-0.18	[-0.74, 0.34]	.52
Shifting (z-stand.)	0.24	[-0.23, 0.70]	.31
Arithmetic fluency (z-stand.)	0.21	[-0.27, 0.69]	.39
Item level predictors			
Problem category switch	-0.24	[-0.52, 0.03]	.09
Mixed-unit problem (vs. small-unit)	-0.88	[-1.21, -0.55]	< .001
Large-unit problem (vs. small-unit)	-1.12	[-1.47, -0.79]	< .001
Problem size (per 10 unit increase)	-0.09	[-0.13, -0.04]	< .001
Random Part	***u***	**95% CI**	**Δ *DIC***
Participant intercept variance	3.86	[2.43, 6.02]	851.2
Item intercept variance	0.11	[0.01, 0.26]	11.1

Note. 95% CI = 95% credible interval; **Δ** DIC = change in Bayesian Deviance Information Criterion, if random intercept dropped from model.

Effects on estimation latencies. Estimation latencies were times between the item appearing on screen and children completing their responses. Therefore, estimation latencies include the time it takes to encode the problem, to select a strategy for the problem, to execute the estimation strategy including adding the rounded numbers and to type in the estimate. If children solved tasks with an inflexible approach not adapting the estimation strategy to the problems but using one dominant strategy, problems were solved on average in 3.50 s (β = 3.50, 95% CI [3.14, 3.86]).

All further analyses were performed only on trials from test blocks which were approached in a flexible manner. Please note that LMMs allow to model effects at the trial level, therefore, effects of how students responded to particular items, for example that on a given trial a child chose mixed-rounding and had switched strategy. To examine estimation latencies in a design with strategy choice (and not strategy execution speeds in a no-choice condition; like Imbo & LeFevre, 2011; Lemaire & Lecacheur, 2002) this feature of including chosen responses as predictors is essential. Additionally, strategy selection effects can and do occur in a choice-design. As demonstrated in the previous section, children were less likely to choose the best strategy on items with larger problem size. Two further features of LMMs are essential to avoid biased estimation latencies. Expected effects on estimation latencies like problem size can be and were included as fixed effects and further unmeasured sources of item difficulty affecting estimation times can be and were modelled as random item intercept variance.

Children needed on average about 5 to 6 s to come up with estimates (compare Table 4 for model parameters, CIs and p-values). Again, there were clear inter-individual differences in estimation latencies. For example, a slow child (1 SD below the mean) needed 1.33 s longer than an average child to solve each estimation problem. However, neither children’s working memory capacity (p = .48) nor their shifting capabilities (p = .18) could explain differences in estimation latencies. In contrast, arithmetic fluency clearly was predictive (p < .001). Children who were 1 SD faster than the mean in arithmetic fluency were 0.81 s faster than an average child in solving each computational estimation problem.

To examine potential switching costs on estimation latencies, trials were classified as strategy switch trial or strategy repetition trials based on each child’s strategic behaviour on the current and the previous trial. Using a different rounding strategy than on the previous trial, that is switching strategy, was not associated with a significant detrimental effect in estimation latencies (p = .38). Hence, no significant switching cost was found.

The strategy used on a trial and the problem size were included as further predictors in the analysis, to examine whether effects of these variables generalize to the current design and to control for the effect of these variables when examining effects of executive functions and switching costs. As expected, children were fastest when solving problems by rounding down: they needed on average 5.19 s. They were significantly slower when using the mixed-rounding strategy (5.62 s; p < .001). Trials solved by rounding-up had the longest estimation latencies with 6.07 s (p < .001). The problem size, defined as the size of the best estimate to an item, had the expected clear effect: estimation latency increased by 0.16 s (p < .001) for each 10-unit increase in problem size. Therefore, a task with 160 as estimate was solved 1.6 s slower than a task with 60 as estimate.

Table 4

LMM examining predictors of estimation latency

Fixed Part	**β**	**95% CI**	***p***
Intercept (reference strategy rounding-down)	5.19	[4.88, 5.49]	< .001
Participant level predictors			
Working memory (z-stand.)	-0.08	[-0.32, 0.14]	.48
Shifting (z-stand.)	0.17	[-0.07, 0.41]	.18
Arithmetic fluency (z-stand.)	0.81	[0.55, 1.07]	< .001
Trial level predictors (participant at item)			
Strategy switch	0.07	[-0.08, 0.21]	.37
Mixed-rounding (vs. rounding-down)	0.43	[0.24, 0.62]	< .001
Rounding-up (vs. rounding-down)	0.88	[0.67, 1.10]	< .001
Item level predictors			
Problem size (per 10 unit increase)	0.16	[0.12, 0.20]	< .001
Random Part	***u***	**95% CI**	**Δ *DIC***
Participant intercept variance	1.04	[0.71, 1.50]	*NA*[1)]
Item intercept variance	0.25	[0.16, 0.37]	254.2
Residual variance	2.77	[2.65, 2.90]	

Note. 95% CI = 95% credible interval; **Δ** DIC = change in Bayesian Deviance Information Criterion if random intercept dropped from model.

1) No meaningful model comparison available as model without random participant intercepts had estimation problems. Both the clearly non-zero participant intercept variance and the estimation problems (if this variance was omitted) indicated that there were definitely between participant differences in estimation latencies.

Discussion

Children with a flexible approach to strategy selection adapted their chosen rounding strategies on average well to the unit sizes of the problems with about 80-90% of trials solved with the best strategy and they solved tasks on average in about 5 s to 6 s. At the same time, significant differences between participants and between estimation problems were present.

However, neither individual differences in the adaptivity of strategy choices nor individual differences in estimation latencies could be explained by children's working memory or shifting capabilities. Only arithmetic fluency predicted how fast children solved estimation problems. Not finding significant relationships between executive functions and best strategy use or estimation latencies is inconsistent with previous findings with the restricted design (Hammerstein et al., 2019; Hodzik & Lemaire, 2011, Lemaire & Lecacheur, 2011). Non-significant results were not due to unreliability of the predictors, as each cognitive capability was measured reliably with multiple tasks. It is also unlikely, that the high proportion of problems solved with the best strategy was the main cause for not finding significant relationships. Firstly, logistic GLMMs (unlike ANOVAs) can properly model proportions approaching 100%. Secondly the large number of 72 problems solved per child allowed to distinguish between children who adapted their strategy choices well and those who adapted them very well to the problem. And thirdly, the participant intercept variance was clear evidence that not all students were at ceiling but that there were reliable differences in strategy adaptivity (and estimation latencies) between participants. As non-significant effects cannot be explained with methodological shortcomings, differences in executive function could have a smaller impact on computational estimation in a more natural setting than previous results from studies restricting mixed-rounding might have suggested. This conclusion is in line with findings by Seethaler & Fuchs (2006) that the predictive effect of working memory on computational estimation was significant, but small in comparison to the impact of arithmetic skills.

Similarly, neither problem category switches affected the likelihood of selecting the best strategy nor switching strategy affected estimation latencies. Switching cost could be smaller in the current design allowing mixed-rounding, if switching to the preferred strategy without strategy restrictions is less demanding on executive functions. Alternatively, switching costs could be less relevant for computational estimation than other factors, irrespective of whether mixed-rounding is allowed or excluded. Firstly, while previous studies with the design restricted to two strategies conclude that switching costs or repetition benefits exist, their result sections revealed a mixed picture (only small switching costs for small-unit but not large-unit problems, Lemaire & Lecacheur, 2010; only repetition benefits for the second of two always repeated prime problems, Lemaire & Brun, 2016; Lemaire & Leclère, 2014). Secondly, switching costs as longer reaction times for task-switch trials compared to task-repetition trials have been demonstrated across a range of *single step tasks* (Monsell, 1996) like categorizing digits, letters, etc. according to one rule (e.g. magnitude categorization of digits) or another rule (e.g. parity categorization of digits; see Kiesel et al., 2010). Compared to these single step tasks, computational estimation is complex with intermediate steps: choosing an approximation strategy, carrying out a certain rounding approach, updating rounded operands in working memory and then doing mental calculation to derive an estimate. If there are switching costs in computational estimation, these costs seem to be small in comparison to other effects of the multiple steps: like the effect of carrying out the chosen rounding strategy or the effect of problem size (see results) which is probably linked to the last step of mental calculation (see LeFevre, Sadesky & Bisanz, 1996).

In contrast to not finding significant effects of executive functions or switching costs, the impact of problem features on strategy choice and on estimation latencies found in studies with a restricted strategy set clearly generalized to the current design. Consistent with previous research (e.g. Lemaire & Brun, 2016; Lemaire & Lecacheur, 2002, 2010), children were somewhat less likely to choose the best strategy for large-unit compared to small-unit problems and were slower when using the rounding-up strategy compared to rounding-down. Additionally, the present study extended previous findings by showing that estimation latencies for mixed-rounding were half-way between rounding-down and rounding-up and that the likelihood of choosing the best strategy for mixed-unit problems was between those for small-unit and large-unit problems. Furthermore, the finding that children need more time to solve problems with large estimates (Hammerstein et al., 2019) generalized to the current study with mixed-rounding allowed. This was expected, as this problem size effect on solution times is one of the most robust findings across different tasks in arithmetic (Zbrodoff & Logan, 2005). Finding these effects supports the notion, that the failure to detect switching costs or effects of executive functions cannot be merely explained by flaws in the present design or a lack of power.

General Discussion

The present study was the first focussing on children's mixed-rounding strategy use in computational estimation. Results showed, that fourth graders, who adjusted their strategy use, clearly preferred mixed-rounding for mixed-unit problems and that unit sums typically were not considered in strategy selection as it would be if small-unit problems like 53 + 74 would be solved by mixed-rounding to obtain a closer estimate. However, being the first study of its kind, it is yet unknown whether these findings replicate in other samples with children of similar age and generalizes to other ages. But given the clear result pattern that was consistent with less detailed results on mixed-rounding in undergraduates (Xu et al., 2014) and consistent with evidence that children rarely use compensation (LeFevre et al., 1993; Lemaire et al., 2000), the present results clearly suggest that mixed-rounding should not be excluded when examining the involvement of executive functions in computational estimation. Firstly, preventing children from using their preferred mixed-rounding approach for mixed-unit problems and asking them to find the second best strategy for which they additionally need to consider unit sums could increase demands on executive functions. Secondly, analyses in the second part of the paper could not show significant effects of working memory and shifting capabilities nor switching costs with the more naturalistic design allowing mixed-rounding. Therefore, previous results with the restricted design could have overestimated the effect of executive functions.

This does not imply that working memory or shifting capabilities were irrelevant for executing computational estimation tasks. Without any working memory resources children would not be able to provide meaningful estimates as they could not hold rounded numbers in working memory. And without the ability to shift between strategies they could not have been classified as using a flexible approach. However, the non-significant or small relationships could indicate that inter-individual differences in executive functions could be less important in determining how well children adapt their strategy use to problem features than some previous studies might have suggested. Gaining an accurate understanding of how relevant differences in executive functions are for computational estimation, has practical ramifications for teaching arithmetic. If there is no link or only a weak link between executive functions and computational estimation in an ecologically valid design, children could become good estimators irrespective of their executive functions. If there, in contrast, is a substantial link, children with lower executive functions would need additional support to become proficient estimators.

All published studies into computational estimation and executive functions (that we are aware of, including this study) asked participants to produce estimates for problems which consisted of two operands with typically two-digits each and which were presented in a visual format without time-pressure. Varying problem features like increasing the number of operands or the number of digits of operands (e.g. LeFevre et al., 1993), like presenting problems in an auditory transient form instead of a stable visual one or like imposing time-pressure could increase task difficulty and manipulate potential demands on executive functions in an ecologically meaningful way without the need to restrict strategy choice. If problems are structured in a way that participants don't have to produce estimates but compare estimates to an anchor, they use a sense of magnitude strategy for certain problems (Ganor-Stern, 2016) which might have lower demands on executive functions. Therefore, only future research with variations in task design can provide a full picture of the potential role that executive functions play for computational estimation.

References

Ai, J., Yang, J., Zhang, T., Si, J., & Liu, Y. (2017). The effect of central executive load on fourth and sixth graders' use of arithmetic strategies. *Psychologica Belgica*, *57*, 154–172. doi:10.5334/pb.360

Browne, W. J. (2017). *MCMC Estimation in MLwiN v3.00*. Centre for Multilevel Modelling, University of Bristol.

Charlton, C., Rasbash, J., Browne, W. J., Healy, M., & Cameron, B. (2018). *MLwiN Version 3.02*. Centre for Multilevel Modelling, University of Bristol.

Dirk, J., & Schmiedek, F. (2016). Fluctuations in elementary school children's working memory performance in the school context. *Journal of Educational Psychology*, *108*, 722–739. doi:10.1037/edu0000076

Dowker, A. (1997). Young children’s addition estimates. *Mathematical Cognition, 3*, 140–153. doi:10.1080/135467997387452

Espy, K. A. (1997). The shape school: Assessing executive function in preschool children. *Developmental Neuropsychology, 13*, 495–499, doi:10.1080/87565649709540690

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

Ganor-Stern, D. (2015). When you don’t have to be exact: Investigating computational estimation skills with a comparison task. *Acta Psychologica*, *154*, 1–9. doi:10.1016/j.actpsy.2014.10.010

Ganor-Stern, D. (2016). Solving math problems approximately: A developmental perspective. *PLoS ONE, 11(5)*. doi:10.1371/journal.pone.0155515

Hammerstein, S., Poloczek, S., Lösche, P., Lemaire, P., & Büttner, G. (2019). Effects of working memory updating on children's arithmetic performance and strategy use: A study in computational estimation. *Journal of Experimental Child Psychology, 184*, 174–191. doi:10.1016/j.jecp.2019.04.003

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods, 50*, 1166–1186. doi:10.3758/s13428-017-0935-1

Hodzik, S., & Lemaire, P. (2011). Inhibition and shifting capacities mediate adults’ age-related differences in strategy selection and repertoire. *Acta Psychologica*, *137*, 335–344. doi:10.1016/j.actpsy.2011.04.002

Hoffman, L., & Rovine, M. J. (2007). Multilevel models for the experimental psychologist: foundations and illustrative examples. *Behavior Research Methods, 39(1),* 101–117. doi:10.3758/BF03192848

Imbo, I., & LeFevre, J. A. (2011). Cultural differences in strategic behavior: A study in computational estimation. *Journal of Experimental Psychology: Learning Memory and Cognition, 37*, 1294–1301. doi:10.1037/a0024070

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59,* 434–446. doi:10.1016/j.jml.2007.11.007

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching-a review. *Psychological Bulletin, 136*, 849–874. doi:10.1037/a0019842

Kray, J., Eber, J., & Karbach, J. (2008). Verbal self-instructions in task switching: a compensatory tool for action-control deficits in childhood and old age? *Developmental Science*, *11*, 223–236. doi:10.1111/j.1467-7687.2008.00673.x

LeFevre, J. A., Greenham, S. L., & Waheed, N. (1993). The Development of Procedural and Conceptual Knowledge in Computational Estimation. *Cognition and Instruction, 11*, 95–132. doi:10.1207/s1532690xci1102_1

LeFevre, J. A., Sadesky, G. S., & Bisanz, J. (1996). Selection of procedures in mental addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning Memory and Cognition, 22*, 216–230. doi:10.1037/0278-7393.22.1.216

Lemaire, P., & Brun, F. (2014). Effects of strategy sequences and response-stimulus intervals on children's strategy selection and strategy execution: A study in computational estimation. *Psychological Research, 78,* 509–519. doi:10.1007/s00426-013-0501-0

Lemaire, P., & Brun, F. (2016). Age-related differences in children's strategy repetition: A study in arithmetic. *Journal of Experimental Child Psychology, 150*, 227–240. doi:10.1016/j.jecp.2016.05.014

Lemaire, P., & Lecacheur, M. (2002). Children's strategies in computational estimation. *Journal of Experimental Child Psychology, 82*, 281–304. doi:10.1016/S0022-0965(02)00107-8

Lemaire, P., & Lecacheur, M. (2010). Strategy switch costs in arithmetic problem solving. *Memory and Cognition, 38*, 322–332. doi:10.3758/MC.38.3.322

Lemaire, P., & Lecacheur, M. (2011). Age-related changes in children's executive functions and strategy selection: A study in computational estimation. *Cognitive Development, 26,* 282–294. doi:10.1016/j.cogdev.2011.01.002

Lemaire, P., Lecacheur, M., & Farioli, F. (2000). Children's strategy use in computational estimation. *Canadian Journal of Experimental Psychology, 54*, 141–148. doi:10.1037/h0087336

Lemaire, P., & Leclère, M. (2014). Strategy repetition in young and older adults: A study in arithmetic. *Developmental Psychology, 50,* 460–468. doi:10.1037/a0033527

Lemaire, P., Luwel, K., & Brun, F. (2017). Does the number of available strategies change how children perform cognitive tasks? Insights from arithmetic. *Journal of Educational and Developmental Psychology, 7*(2), 43–50. doi:10.5539/jedp.v7n2p43

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*, 49–100. doi:10.1006/cogp.1999.0734

Monsell, S. (1996). Control of mental processes. In V. Bruce (Ed.), *Unsolved mysteries of the mind* (pp. 93–148). Hove, England: Erlbaum.

Reys, R. E., Rybolt, J. F., Bestgen, B. J., & Wyatt, J. W. (1982). Processes used by good computational estimators. *Journal for Research in Mathematics Education, 13*, 183–201.

Seethaler, P. M., & Fuchs, L. S. (2006). The cognitive correlates of computational estimation skill among third-grade students. *Learning Disabilities Research and Practice, 21*, 233–243. doi:10.1111/j.1540-5826.2006.00220.x

Siegler, R. S., & Booth, J. L. (2005). Development of numerical estimation. A Review. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 197–212). New York and Hove: Psychology Press. doi:.1037//0033-2909.I26.1.78

Sowder, J. T., & Wheeler, M. M. (1989). The development of concepts and strategies used in computational estimation. *Journal for Research in Mathematics Education, 20*, 130–146. doi:10.2307/749278

Star, J. R., & Rittle-Johnson, B. (2009). It pays to compare: An experimental study on computational estimation. *Journal of Experimental Child Psychology, 102*, 408–426. doi:10.1016/j.jecp.2008.11.004

Star, J. R., Rittle-Johnson, B., Lynch, K., & Perova, N. (2009). The role of prior knowledge in the development of strategy flexibility: The case of computational estimation. *ZDM - International Journal on Mathematics Education, 41*, 569–579. doi:10.1007/s11858-009-0181-9

Xu, C., Wells, E., LeFevre, J. A., & Imbo, I. (2014). Strategic flexibility in computational estimation for Chinese- and Canadian-educated adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 1481–1497. doi:10.1037/a0037346

Zbrodoff, N. J., & Logan, G. D. (2005). What everyone finds. The problem-size effect. In Campbell; J. I. D (Ed.), *Handbook of mathematical cognition* (pp. 331–346). New York and Hove: Psychology Press.

Zhang, Z., Parker, R. M. A., Charlton, C. M. J., Leckie, G., & Browne, W. J. (2016). R2MLwiN: A Package to Run MLwiN from within R. *Journal of Statistical Software, 72*(10), 1–43. doi:10.18637/jss.v072.i10

Supplementary Material

A Computational Estimation Experiment

Further constraints in item construction

As in previous research on computational estimation (e.g. Lemaire & Brun, 2016; Lemaire & Lecacheur, 2011) estimation problems were additionally created with the following constraints: (1) operands never included 0 or 5 as unit digit, (2) no pair of operands included equal unit digits (e.g., never 32 + 62), (3) no pair of operands would result in equal decade digits when using the rounding strategy indicated by the main category (e.g., never 39 + 42, because it would lead to 40 + 40), (4) no trials with reverse order of the operands were presented (e.g., not 68 + 24 in one trial and 24 + 68 in another) and (5) no rounded operand was equal to 0, 10 or 100 if rounded with the best strategy (i.e., range of operands from 16 to 94). Therefore, problem sizes of the estimation items ranged from estimates of 50 to 170.

Table S1

Exact pseudorandom sequences of computational estimation problems

Testblock 1a		Testblock 1b		Testblock 2a		Testblock 2b	
Problem	**Category**	**Problem**	**Category**	**Problem**	**Category**	**Problem**	**Category**
48 + 72	mixed5	42 + 18	mixed5	43 + 61	small1	29 + 42	mixed6
22 + 57	mixed4	43 + 48	mixed6	68 + 16	large7	59 + 87	large9
62 + 54	small3	18 + 29	large9	39 + 48	large9	79 + 16	large8
84 + 71	small2	78 + 37	large8	94 + 32	small3	53 + 94	small3
56 + 28	large7	54 + 93	small3	47 + 19	large9	22 + 74	small3
77 + 29	large9	81 + 24	small2	68 + 59	large9	43 + 22	small2
47 + 28	large8	27 + 73	mixed5	72 + 93	small2	17 + 58	large8
93 + 77	mixed5	41 + 76	mixed4	54 + 33	small3	87 + 76	large7
32 + 83	small2	93 + 31	small1	62 + 81	small1	46 + 84	mixed5
19 + 76	large8	41 + 63	small1	64 + 28	mixed6	16 + 93	mixed4
38 + 86	large7	63 + 24	small3	91 + 47	mixed4	92 + 59	mixed6
88 + 43	mixed6	27 + 36	large7	19 + 43	mixed6	41 + 32	small1
23 + 52	small2	69 + 88	large9	78 + 47	large8	16 + 37	large7
34 + 62	small3	76 + 57	large7	46 + 58	large7	36 + 69	large8
81 + 92	small1	87 + 52	mixed4	27 + 51	mixed4	73 + 26	mixed4
66 + 21	mixed4	39 + 34	mixed6	54 + 36	mixed5	28 + 94	mixed6
32 + 21	small1	59 + 37	large9	71 + 43	small1	31 + 74	small2
74 + 79	mixed6	56 + 69	large8	84 + 41	small2	93 + 69	mixed6

Note. Items were presented to all participants in the same pseudorandom, but carefully balanced order (in each testblock from top to bottom).

Sequence of trials

As we planned to analyse inter-individual differences, all children were presented with the same 72 items in the same pseudorandom order. As we additionally were interested in comparing strategy selection and estimation latencies between categories of problems, the pool of estimation problems as well as the pseudorandom test order were carefully constructed to reduce the impact of potential confounding variables. The 72 items were distributed onto the four test blocks so that test blocks contained six problems of each main category and that blocks were well matched on the other dimensions listed above. Within each test block, problems were put into random order. This random sequence was adjusted so that there were about 50%

of trials in which the same main task category (small, mixed, large) was repeated. If children always or mostly used the rounding strategy suggested by the main category, the number of strategy repetitions and switches was also balanced or nearly balanced. Additionally, the split between suggested strategy repetition and suggested strategy switches was similar for all nine subcategories. Finally, after rounding no two consecutive trials could be the same (e.g. 61 + 43 could not be followed by 57 + 41 because both would result in 60 + 40; for the exact sequence of items see Table S1).

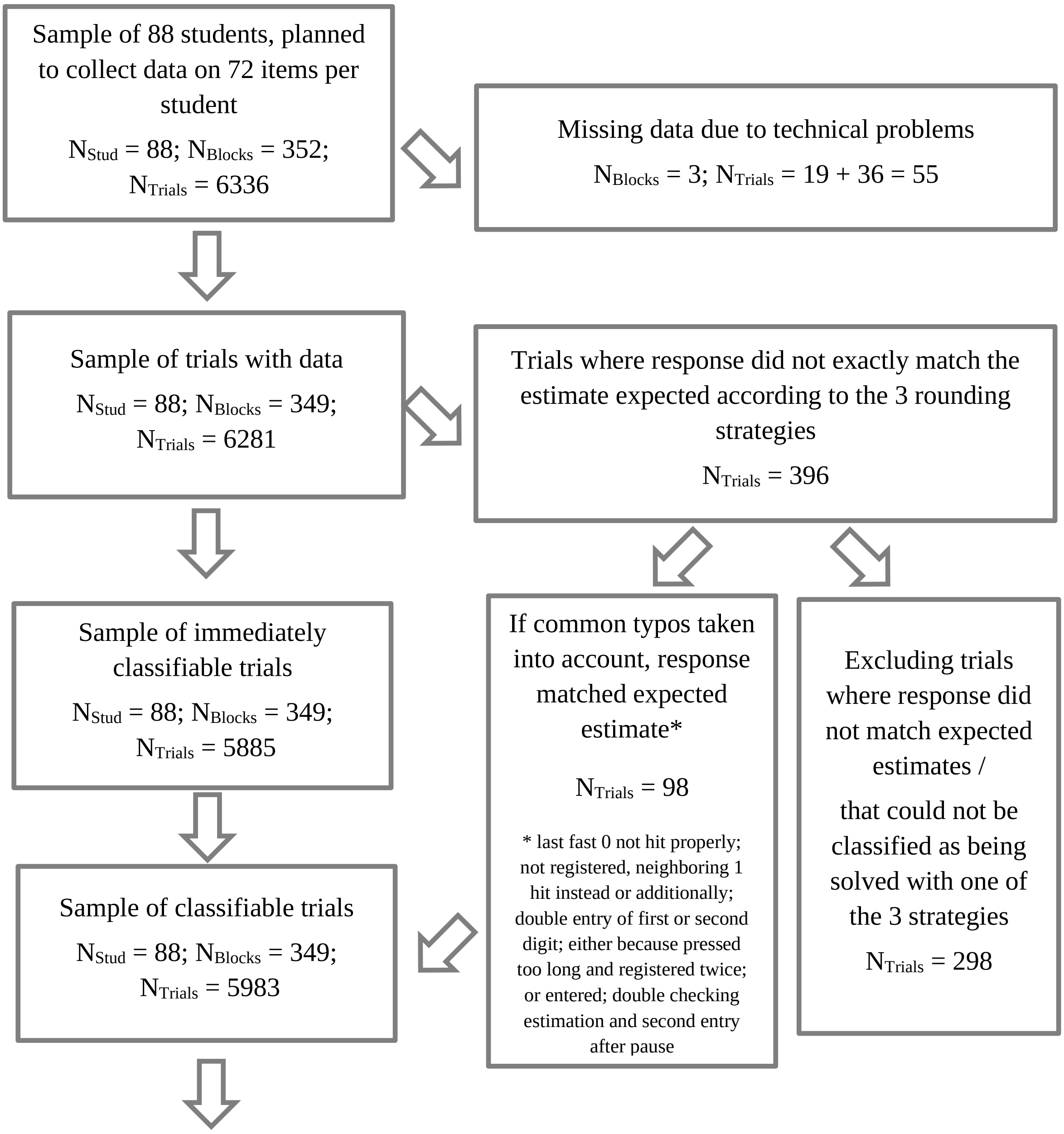
Inferring strategy use

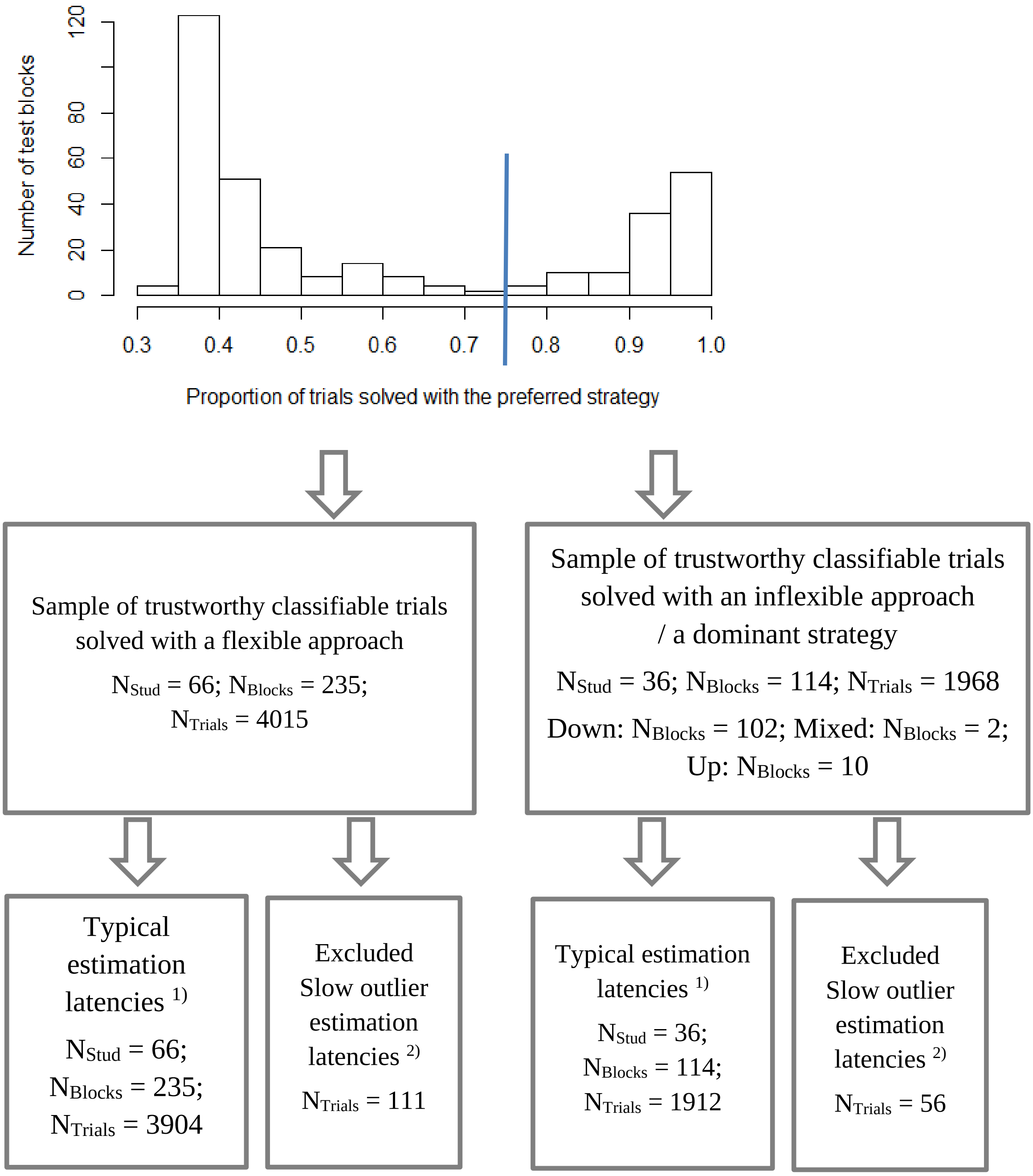
The rounding strategy selected by a child for a given trial was inferred from the estimate for that trial. If for example a child typed in 70 as response to 22 + 57, the strategy was classified as rounding-down; if the response was 80 as mixed-rounding; and if the response was 90 as rounding-up. In 93.7% of all 6281 trials the response directly matched one of expected estimates. Of the remaining 396 responses 98 very likely were due to common typos like not properly hitting the last 0 digit (entries like 71 or 710 or 7 for the example above) or having either the first digit or the second digit logged twice (entries like 770 or 177 for problems in which 70 or 170 were plausible estimates). The syntax to classify responses was extended so that also trials with common typos were classified, resulting in 95.2% of trials having a strategy classification.

To assess whether determining strategy choice based on the estimates would likely lead to misclassifications, we examined children's calculation errors during the pure calculation task of two-digit additions (e.g., 40 + 70; data gathered to estimate children's addition speed, see Assessment of cognitive abilities - Arithmetic fluency). Within a total of 1056 trials across 88 participants, 93.9% of entered responses were correct, further 2.7% contained an error that would not lead to a strategy-misclassification (e.g. first digit not registered, typo or calculation error not close to correct result). There were 1.6% +10-errors (i.e., answers with 10 units above the correct result; e.g. 120 instead of 110) and 0.9% -10-errors (i.e., answers with 10 units below the correct result). These errors would lead to classification errors, if they were paired with a mixed-unit problem, and would lead to classification errors with a 50% chance when paired with small-unit or large-unit problems. +20-errors and -20-errors were very rare (0.3% and

0.5%, respectively) and just one in three of those errors would lead to a misclassification. Therefore, inferring strategy use from the given responses in the computational estimation task should only be associated with a small error.

Flowchart inclusion/exclusion of trials





Note. To reduce the impact of outliers on the analyses of estimation latencies, only latencies within the boundaries of +/-2.5 standard deviations (computed for each individual) around the individuals mean latency were included (considered typical estimation latencies). No fast outliers were present. Slow outliers were present on 2.7% of trials and were excluded.

1) RTs or estimation latencies that were within +/- 2.5 (individual) SDs from the individual’s mean
2) RTs or estimation latencies that were slower than + 2.5 (individual) SDs above the individual’s mean

B Further Details on the Assessment of Cognitive Capabilities

Working memory

Working memory was assessed by three subtests. First, participants were administered a *Spatial Updating* task (Dirk & Schmiedek, 2016). Children were presented with coloured, fictitious creatures at different positions on a 4-by-4 grid. Participants were asked to remember and update the positions of two creatures on eight trials and of three creatures on further eight trials. The starting positions of all creatures were presented for 3000ms and then were masked. Then, an arrow matching the colour of one creature was presented in the centre of the grid for 2500ms. The arrow indicated in which direction the creature with the matching colour covertly moved. The position of this creature had to be updated as the arrow only indicated the direction of movement and not the new covert position. On trials with two creatures, in total three updating operations were required, i.e. one creature moved twice and the other moved once; on trials with three creatures, four arrows were shown to prompt updating. At the end, children were asked to reproduce the final positions of the creatures on an empty grid.

During 16 test trials of the *Colour Updating* task (adapted from Pictorial Updating tasks; e.g., Lee, Ng, Bull, Pe, & Ho, 2011), participants were visually and verbally presented with sequences of colours, one at a time. Stimuli consisted of eight colours with monosyllabic names in German. Each colour was presented for 1500ms before it was covered and 500ms later the next colour appeared. Presentation stopped after 4 to 8 colours. The task required updating as children were asked to remember either the last two or three colours (8 trials each).

Furthermore, we used a *Picture Span Backward* task to assess children's working memory functions. On 20 trials, children were presented with a sequence of pictures, each one for 1500ms with inter-stimulus-intervals of 500ms. Simultaneously each picture name was played. The stimuli set consisted of nine black-and-white line drawings, each representing a German, monosyllabic, highly imageable word with a low age of acquisition. At the end of each trial consisting 2 to 5 items (five trials per list length), participants were asked to reproduce the pictures in reverse order on a response array with nine pictures. This task set-up was chosen to use the non-verbal response format of a picture span task (e.g., Henry & Winfield, 2010) while preserving as many features as possible of a verbal backward span task (e.g., St Clair-Thompson & Gathercole, 2007).

The proportion of correctly remembered items within each of the three tasks served as indicators for children's working memory capacity. To estimate reliability, for each task four parcels – each containing one quarter of the items – were built. The internal consistency of the working memory tasks was satisfactory to good, with α = .88 for the Spatial Updating task, α = .76 for the Colour Updating task and α = .73 for the Picture Span Backward task (see Table S2). The internal consistency across all 12 working memory parcels was α = .85 (see Table S3). The average performance across all *z*-standardised parcels served as indicator or each child's working memory capacity.

Table S2

Reliability statistics of each working memory task

	Colour Updating	**Picture Backwards**	**Spatial Updating**
Number of Items (trials*items per trial)	8*(2+3) = 40	5*(2+3+4+5) = 60	8*(2+3) = 40
Cronbach´s α	.76	.73	.88
Overall Score	coup_pc	piba_pc	spup_pc
Sample	87 [1)]	87 [2)]	80 [3)]

1) Data file missing for 1 participant.
2) 1 participant removed from analysis. Performance set to missing as picture sequences reproduced in perfect (at least with 2 words) forward order instead of backward order. Therefore, uncharacteristically low score not representative of working memory capacity. Failing to remove outlier would produce α = .81.
3) Accuracy data for 8 participants not properly logged.

Table S3

Reliability of the working memory composite score

Working Memory – Proportion Correct Scores	**M**	**SD**	**r_{it}**	**α**
Colour Updating – Parcel 1	.72	.21	.67	
Colour Updating – Parcel 2	.60	.20	.25	
Colour Updating – Parcel 3	.67	.24	.40	
Colour Updating – Parcel 4	.63	.23	.49	
Picture backwards – Parcel 1	.73	.16	.58	
Picture backwards – Parcel 2	.74	.14	.39	
Picture backwards – Parcel 3	.73	.14	.30	
Picture backwards – Parcel 4	.69	.16	.54	
Spatial Updating – Parcel 1	.62	.28	.61	
Spatial Updating – Parcel 2	.62	.27	.63	
Spatial Updating – Parcel 3	.65	.25	.75	
Spatial Updating – Parcel 4	.65	.27	.61	
Working Memory Composite (Mean of z-scores)				**.85**

Shifting

We intended to use three subtests to measure children's shifting functions. First, children had to solve an *Animal Colour* task (Kray, Eber, & Karbach, 2008). Stimuli consisted of pictures of 14 animals that can either fly or swim, each in a coloured and a black-and-white version. In the first single task block, participants had to make a quick decision whether a presented animal can swim or fly. In the second single task block, they had to decide whether the picture was coloured or black-and-white. Then in the two mixed blocks, a verbal cue on each trial indicated whether children had to sort the picture according to the type of animal or the colour. Each block comprised 25 trials.

In the *Colour Shape* task (Espy, 1997), participants were told to sort pictures of smileys. In the colour task block, participants had to indicate whether the presented smiley was blue or green. In the shape task block, they had to decide whether the smiley had either a circular or a square shape. In the two mixed colour shape blocks, the direction of each smiley´s arms as visual cue indicated whether children had to sort it regarding the criterion of either colour or shape. Again, each block consisted of 25 test trials.

Furthermore, we adapted the *Trail Making Test* (Reitan, 1992) with the purpose of avoiding a numeric component in the assessment of executive functions. Thus, instead of a sequence of numbers we used a sequence of black-and-white drawings in which a snowman was built. Due to reconstruction errors in the snowman sequences and a lack of reliability (α = .27), this task was excluded from further analyses.

As common practice, only reaction times (RTs) for correct trials were considered. Additionally, fast guessing RTs below 200ms were excluded and RTs were ln-transformed to reduce skewness and the impact of slow RTs on the average. The differences between the RTs in the mixed task blocks and the single task blocks were taken as shifting indicators; thus, higher scores indicate larger switching cost or lower shifting capabilities. Note that because of the requirements in the mixed task blocks, the shifting indicator comprises not only switching but also mixing costs.

The internal consistency of these difference scores (based on four parcels per task) was α = .61 for the Animal Colour task and α = .81 for the Colour Shape task (see Table S4). The internal consistency of the shifting composite-score, that was computed as the mean of the eight *z*-standardised difference score parcels, was α = .73 (see Table S5).

Table S4

Reliability statistics of each shifting task

	Animal Colour	**Colour Shape**	**Trail Making**
Number of Items	2*(24+24)=48	2*(24+24)=48	2*4*(8+8)=128
Cronbach´s α Difference Score	.61	.81	.27
Cronbach´s α RTs in Single	.88	.93	.92
Cronbach´s α RTs in Mixed	.91	.93	.88
Correlation RTs Mixed - Single	.78	.63	.63
Overall Score	anco_lnRTacc_diff	cosh_lnRTacc_diff	not computed
Sample	86 / 87 [1)]	85 / 88 [2)]	34 / 45 / 69 [3)]

1) 1 participant failed to reach 75% accuracy in both mixed blocks; therefore, no overall difference score available. 1 participant failed to reach 75% accuracy in one of the mixed blocks; for this participant only partial data entered the reliability analysis and the calculation of the overall score as average.

2) In one of the mixed blocks, 3 participants failed to reach 75% accuracy; for these participant only partial data entered the reliability analysis; their overall scores are based on the available difference scores for the other test block.

3) Varying sample size because data for the single letter condition in several children not logged even though task completed and because several children failed to reach 75% accuracy in the single snowman and the mixed snowman condition. Low reliability (potentially caused or aggravated by difficulties in reconstructing the snowman sequences) lead to the decision to not include task performance as a shifting indicator.

Table S5

Reliability of the shifting cost composite score

Shifting Costs – Differences in ln(RTs)	**M**	**SD**	**r_{it}**	**α**
Animal Colour – Animal Difference Parcel 1	0.41	0.19	.30	
Animal Colour – Animal Difference Parcel 2	0.42	0.21	.29	
Animal Colour – Colour Difference Parcel 1	0.68	0.20	.36	
Animal Colour – Colour Difference Parcel 2	0.61	0.18	.38	
Colour Shape – Shape Difference Parcel 1	0.56	0.18	.52	
Colour Shape – Shape Difference Parcel 2	0.51	0.20	.55	
Colour Shape – Colour Difference Parcel 1	0.70	0.21	.51	
Colour Shape – Colour Difference Parcel 2	0.61	0.18	.49	
Shifting Cost Composite (Mean of z-scores)				**.73**

Inhibition

Inhibition was assessed by four measures: A *Real Animal Size* test (Catale & Meulemans, 2009), an *Object Inhibition* task (e.g., van der Sluis, de Jong, & van der Leij, 2004), a *Flanker* task (Rueda et al., 2004) and a standard *Colour Stroop* task (Stroop, 1935) were adapted for computerized administration without verbal responses.

Each task consisted of a practice and three test blocks with 116 trials in total. To obtain scores that are comparable across participants and across test blocks pseudorandom trial sequences were constructed. Each test block consisted of one start trial, 18 inhibition and 18 control trials. Inhibition and control trials were matched on response key, response repetition (i.e., number of trials responding with the same hand as on the previous trial vs. switching response hand) and condition repetition (i.e., number of trials in which the condition - inhibition or control - was the same as on the previous trial). Although trial sequences in this study were carefully balanced and although the score for each task was based on 96 trials in 3 test blocks, thus, 3 difference scores, the reliability estimates for all inhibition tasks were unacceptably low (ranging from α = -.11 to α = .49; see Table S6). This was despite high internal consistency scores of the reaction times for the inhibition trials (ranging from α = .85 to α = .93) and for the control trials (ranging from α = .85 to α = .94) and despite finding slower reaction times for inhibition trials than control trials at group level. Problems regarding the reliability of inhibition scores as individual difference measures are not specific to just this study, but are starting to be acknowledged and reported more widely in research (compare Hedge, Powell, & Sumner, 2018). Therefore, the inhibition tasks were excluded from further analyses.

Table S6

Reliability statistics of each shifting task

	Animal Stroop	**Colour Stroop**	**Flanker**	**Object Inhibition**
Number of Items	3*(16+16)=96	3*(16+16)=96	3*(16+16)=96	3*(16+16)=96
Cronbach´s α Difference	-.01	-.11	.02	.49
Cronbach´s α RTs in Control Trials	.85	.88	.92	.89
Cronbach´s α RTs in Inhibition Trials	.85	.91	.93	.90
Correlation RTs Inhibition - Control	.91	.95	.96	.91
Overall Score	not computed	not computed	not computed	not computed

Correlations and factorial structure of executive function tasks

Table S7 displays the correlations between the five reliable executive function tasks (scores are the mean of the 4 z-standardized parcel scores of a tasks). The three working memory tasks were correlated to each other and the two shifting tasks were correlated. No significant correlations between working memory and shifting tasks were observed. The pattern of correlations is consistent with our plan, to integrate the score of the five tasks into a working memory and a shifting indicator as the tasks measuring one construct were correlated. At the same time, the correlations were considerably lower than the reliabilities/internal consistencies of each task. This suggests that apart from generic and stable working memory or shifting variance each task measured task and / or situation specific variance, highlighting the need to assess each executive function with multiple tasks.

Table S7

Correlations between all reliable executive function tasks (with 95% confidence intervals)

Variable	Colour Updating	Picture Backwards	Spatial Updating	Animal Colour
Picture Backwards	.37** [.17, .54]			
Spatial Updating	.35** [.14, .53]	.48** [.29, .63]		
Animal Colour	.01 [-.20, .22]	.06 [-.15, .27]	.03 [-.19, .25]	
Colour Shape	-.03 [-.24, .18]	.12 [-.10, .32]	-.07 [-.29, .15]	.24* [.04, .43]

Note. Values in square brackets indicate the 95% confidence interval for each correlation; * indicates $p < .05$. ** indicates $p < .01$.

To determine the factorial structure of the scores in the different executive function task parcels, principal axes factor analyses (extraction method 'OLS', oblique rotation 'oblimin') were performed with the "psych" package in R. Both a 5-factor-solution and a 2-factor-solution were done and are reported; the 5-factor-solution because the parallel analysis suggested five factors and because parcels belonging to 5 tasks were entered into the factor analyses; the 2-factor-solution because we would expect two higher order factors for working memory and for shifting.

Results revealed that in the 5-factor-solution the main loading (highest loading) of all parcels were on the factor representing the corresponding tasks and that in the 2-factor-solution the main loadings of all working memory parcel were on factor 1, while the main loadings of the shifting parcels were on factor 2 (compare Table S8). Therefore, integrating the information of all working memory parcels into one working memory indicator and combining all shifting parcels into one shifting indicator is theoretically and empirically justified.

Table S8

Loading patterns of executive function parcels after principal axes factor analyses with oblique rotation

		5-Factor-Solution					2-Factor-Solution	
		Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 1	Factor 2
Colour Updating	Parcel 1			.72			.67	
	Parcel 2			.61			.25	
	Parcel 3			.71			.40	
	Parcel 4			.61			.50	
Picture Backwards	Parcel 1	.24			.67		.63	
	Parcel 2	.25	.21		.51		.45	.24
	Parcel 3			.34	.47		.33	.24
	Parcel 4	.25			.49		.61	
Spatial Updating	Parcel 1	.87					.71	
	Parcel 2	.84					.71	
	Parcel 3	.73		.23			.82	
	Parcel 4	.64					.68	
Animal Colour	Parcel 1				.31	.49		.28
	Parcel 2	-.24			.21	.49		.26
	Parcel 3					.58		.31
	Parcel 4					.58		.30
Colour Shape	Parcel 1		.77					.71
	Parcel 2		.71					.75
	Parcel 3		.76					.72
	Parcel 4		.63					.63

Note. Only loadings >= .20 are displayed.

Arithmetic fluency

We did not administer a standardized arithmetic performance test but decided to assess the (sub)skills needed to solve computational estimation tasks: to decide, whether a number should be rounded up or down, to carry out rounding and to add numbers. The range of numbers was comparable to the range of numbers in the computational estimation task (for the complete list of items see Table S9). As we expected the accuracy to be high on those tasks, the response times were taken as indicator of inter-individual differences in arithmetic skills. For all tasks, children received practice trials and were instructed to provide the correct answer as fast as possible.

The *Speed of Rounding Decisions* task required children to indicate whether 25 two-digit numbers should be rounded down (e.g., 43) or up (e.g., 68). To respond as fast as possible and to reduce the measurement error, children were instructed to put the right index finger on the up-arrow und the left index finger on the down-arrow.

Additionally, children's *Speed of Cued Rounding* was assessed with 25 test trials. Two-digit numbers appeared on the screen while arrows indicated in which direction to round. The cued direction was always the one that led to the correct rounding result. Participants were asked to enter the correct two-digit response as fast as possible.

Children's *Addition Speed* was operationalized by their performance on 24 visually and verbally presented one- and two-digit addition problems; half of them with and half without carry over ten or one hundred. Operands in the two-digit additions only included zero as unit digit (e.g., 30 + 60).

In all arithmetic tasks, the number or the numbers of the trial were presented in black (Font: Arial, Font size: 150) in the centre of the screen. RTs of incorrect trials and fast guessing RTs below 200ms were excluded. For each arithmetic task, ln-transformed RTs were averaged to yield subtest parcels. The internal consistencies of the three arithmetic tasks (based on four parcels per task) were all very high; $\alpha = .92$ for Addition Speed, $\alpha = .94$ for Speed of Rounding Decision and $\alpha = .93$ for Speed of Cued Rounding (see Table S10). The arithmetic composite score was computed as the average of the *z*-standardised parcels with an internal consistency of $\alpha = .92$ (see Table S11).

In order to not influence rounding strategy use, the tasks Speed of Rounding Decision and Speed of Cued Rounding were presented after the computational estimation tasks.

Table S9

Items of the three arithmetic fluency tasks

Rounding Decision		**Rounding Speed**		**Addition Speed**			
Practice	**Test**	**Practice**	**Test**	**Practice**	**Test**	**Practice**	**Test**
44	16	26	59	3 + 9	8 + 9	60 + 20	20 + 50
61	61	83	36	5 + 2	4 + 6	30 + 80	90 + 80
16	36	46	93	2 + 8	6 + 5	50 + 40	60 + 40
88	68	27	52	9 + 4	7 + 8	40 + 80	60 + 80
26	84	41	84	5 + 7	4 + 5	70 + 60	70 + 50
41	91	58	19	5 + 3	8 + 6	20 + 30	40 + 20
74	42	51	88	3 + 6	2 + 6	50 + 60	40 + 90
77	23	74	62	2 + 4	3 + 2	80 + 20	60 + 30
49	29	32	42	9 + 7	8 + 4	20 + 70	30 + 50
93	87	93	24	8 + 3	3 + 4	70 + 90	30 + 70
53	63	68	71	5 + 9	6 + 7	40 + 30	80 + 70
68	54	16	28	7 + 2	7 + 3	90 + 50	90 + 30
32	82	79	34				
	27		67				
	31		53				
	66		88				
	49		76				
	56		21				
	43		69				
	79		47				
	72		29				
	24		63				
	18		81				
	87		56				
	48		37				

Table S10

Reliability statistics of each arithmetic fluency task

	Rounding Decision Speed	**Rounding (Cued) Speed**	**Addition Speed**
Number of Items	24	24	24
Cronbach´s α	.94	.93	.92
Overall Score	rode_ lnRTacc_m	roun_ lnRTacc_m	addi_lnRTacc_m
Sample	85 [1)]	88	88

1) All children had complete data. However, reaction times of 3 children were excluded because they failed to reach the 75% threshold for valid reaction times. One of them had less than 25% errors (pinc = .23). However, 2 correct responses were excluded because of RTs < 200ms. These rapid guessing RTs were excluded (as all RTs below 200ms were). Therefore, the child did not reach 18 valid RTs.

Table S11

Reliability of the arithmetic fluency composite score

Arithmetic Speed in ln(RTs)	**M**	**SD**	**r_{it}**	**α**
Rounding Decision – Parcel 1	-0.22	0.25	.58	
Rounding Decision – Parcel 2	-0.22	0.23	.63	
Rounding Decision – Parcel 3	-0.28	0.23	.66	
Rounding Decision – Parcel 4	-0.25	0.25	.62	
Rounding Speed – Parcel 1	0.75	0.23	.73	
Rounding Speed – Parcel 2	0.78	0.25	.76	
Rounding Speed – Parcel 3	0.76	0.23	.72	
Rounding Speed – Parcel 4	0.79	0.26	.69	
Addition Speed – Parcel 1	1.06	0.28	.66	
Addition Speed – Parcel 2	.96	0.25	.62	
Addition Speed – Parcel 3	1.02	0.24	.64	
Addition Speed – Parcel 4	0.85	0.22	.70	
Arithmetic Fluency Composite (Mean of z-scores)				**.92**

C Further Details on the Results Section

Dummy coding to compare main task categories and subcategories

The set of eight predictors for the comparison between nine types of task was constructed in a way to be able to compare strategy use at the same time between main task categories (small vs. mixed vs. large unit problems) and within main task categories but between subcategories (small1 vs. small2 vs. small3 | mixed4 vs. mixed5 vs. mixed6 | large7 vs. large8 vs. large9). To achieve this the predictors were set-up as demonstrated in the table on the left.

The first two predictors ensure, that all mixed-unit problems are contrasted with all small-unit problems as main reference category, and likewise large-unit problems are contrasted with small-unit problems.

To compare the three subcategories within main categories, for each main category two subcategory predictors were created. E.g. both small2 and small3 problems (both can be either solved by rounding-down or mixed-rounding) are contrasted to the reference category small1 (twice 0 as dummy predictor; reference category as rounding-down unambiguously best strategy).

For each of the nine subcategories a unique pattern of 8 predictor values exists. All predictors are independent; none can be expressed as a function of the other predictors.

			Main category predictors		Subcategory predictors					
	Main categories	Sub-categories	Mixed (vs. small)	Large (vs. small)	Small2 (vs. small1)	Small3 (vs. small1)	Mixed1 (vs. mixed2)	Mixed3 (vs. mixed2)	Large1 (vs. large3)	Large2 (vs. large3)
Estimation problems	small	small1	0	0	0	0	0	0	0	0
		small2	0	0	1	0	0	0	0	0
		small3	0	0	0	1	0	0	0	0
	mixed	mixed1	1	0	0	0	1	0	0	0
		mixed2	1	0	0	0	0	0	0	0
		mixed3	1	0	0	0	0	1	0	0
	large	large1	0	1	0	0	0	0	1	0
		large2	0	1	0	0	0	0	0	1
		large3	0	1	0	0	0	0	0	0

Calculating predicted probabilities with 95% credible intervals

The predicted probability of choosing a certain strategy for item *i* by participant *j* is

$$\pi_{ij} = \frac{\exp\,(\beta_0 + \beta_1 x_{1ij} + \cdots + \beta_k x_{kij} + u_i + u_j)}{1 + \exp\,(\beta_0 + \beta_1 x_{1ij} + \cdots + \beta_k x_{kij} + u_i + u_j)}$$

where, β_0, β_1 up to β_k for the k^{th} predictor are replaced by the estimates from the fitted GLMM, u_i by the random item intercept residuals and u_j by the random participant intercept residuals.

The predicted probability, that an average person ($u_j = 0$) will choose rounding-down when giving an estimate for an average item ($u_i = 0$) that belongs to the subcategory *small1* (all predictors x_1 to $x_k = 0$) would only include the intercept (see Table 1 in the paper); so, it would be $\exp(3.20)/(1 + \exp(3.20)) = 0.961$.

The formula for the predicted probability for an average person and an average item belonging to the subcategory *mixed4* would additionally include the estimates for main category *mixed*-unit and the subcategory *mixed4*: $\exp(3.20 - 5.70 + 0.19)/(1 + \exp(3.20 - 5.70 + 0.19)) = 0.090$.

To estimate the mean predicted probability for our population of participants and items it is not sufficient to just take the values estimated for an average person and item. Due to the non-linear inverse-logit transformation the mean predicted probability is not equal to the predicted probability of an average participant, the median predicted probability (see Steele, 2009). Mean and median predicted probabilities can be close if probabilities are in the range of 0.20 to 0.80 and if random intercept variance is low. But given the predicted probabilities were more extreme in this study and given that there was considerable random participant variance, mean predicted probabilities with a population-averaged interpretation were calculated. To do so, for each of the 9 task-subcategories 1000 predicted probabilities were calculated, each with a different value for u_i and u_j. These values were drawn from a random distribution with $M = 0$ and variances corresponding to the random item and the random participant variance of the model (compare Steele, 2009). Taking the mean of the 1000 predicted probabilities yields the probability with the population-averaged interpretation.

To provide 95% credible intervals around the mean, this process of simulating 1000 predictions and computing the mean was not performed once with the values provided in Table 1, but 5000 times with the value combinations of the 5000 stored MCMC iterations.

Computing the mean of the 5000 mean predicted probabilities per subcategory and extracting the 2.5th and the 97.5th percentile of these 5000 predicted probabilities resulted in the reported mean predicted probabilities with 95% credible intervals.

We are grateful to Prof. Christopher Jarrold (University of Bristol) for his valuable comments on our thoughts on computing credible intervals for predicted probabilities.

References

Catale, C., & Meulemans, T. (2009). The real animal size test (RAST): A new measure of inhibitory control for young children. *European Journal of Psychological Assessment*, *25*, 83–91. doi:10.1027/1015-5759.25.2.83

Dirk, J., & Schmiedek, F. (2016). Fluctuations in elementary school children's working memory performance in the school context. *Journal of Educational Psychology*, *108*, 722–739. doi:10.1037/edu0000076

Espy, K. A. (1997). The shape school: Assessing executive function in preschool children. *Developmental Neuropsychology, 13*, 495–499, doi:10.1080/87565649709540690

Henry, L., & Winfield, J. (2010). Working memory and educational achievement in children with intellectual disabilities. *Journal of Intellectual Disability Research*, *54,* 354–365. doi:10.1111/j.1365-2788.2010.01264.x

Kray, J., Eber, J., & Karbach, J. (2008). Verbal self-instructions in task switching: a compensatory tool for action-control deficits in childhood and old age? *Developmental science*, *11*, 223–236. doi:10.1111/j.1467-7687.2008.00673.x

Lee, K., Ng, S. F., Bull, R., Pe, M. L., & Ho, R. H. M. (2011). Are patterns important? An investigation of the relationships between proficiencies in patterns, computation, executive functioning, and algebraic word problems. *Journal of Educational Psychology, 103*, 269–281. doi:10.1037/a0023068

Reitan, R. M. (1992). *Trail making test. Manual for administration and scoring*. Tucson, AZ: Reitan Neuropsychological Laboratory.

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., & Posner, M. I. (2004). Development of attentional networks in childhood. *Neuropsychologia*, *42*, 1029-1040. doi:10.1016/j.neuropsychologia.2003.12.012

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology, 59*, 745–759. doi:10.1080/17470210500162854

Steele, F. (2009). Module 9: Single-level and multilevel models for ordinal responses concepts. In Centre for Multilevel Modelling (Eds.). LEMMA (Learning environment for multilevel methods and applications. University of Bristol. Available at: https://www.cmm.bris.ac.uk/lemma/

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662. doi:10.1037/h0054651

van der Sluis, S., de Jong, P. F., & van der Leij, A. (2004). Inhibition and shifting in children with learning deficits in arithmetic and reading. *Journal of experimental child psychology, 87*, 239–266. doi:10.1016/j.jecp.2003.12.002

6.3 Study 3: Two Versus Three Strategies in Strategy Selection

Hammerstein, S., Poloczek, S., Lösche, P., & Büttner, G. (2020). Two versus three available strategies in children's strategy selection in a computational estimation task. *Manuscript submitted for publication.*

Two versus three available strategies in children's strategy selection in a computational estimation task

Brief Report

Svenja Hammerstein[1,2]*, Sebastian Poloczek[1,2], Patrick Lösche[1,2], &
Gerhard Büttner[1,2]

[1]Goethe Universität Frankfurt am Main, Germany

[2]Centre for Individual Development and Adaptive Education of Children at Risk (IDeA), Frankfurt am Main, Germany

*Corresponding author at: Goethe Universität Frankfurt am Main, Theodor-W.-Adorno-Platz 6, 60629 Frankfurt am Main; Tel.: +49 69 798 35377; E-mail address: hammerstein@psych.uni-frankfurt.de (S. Hammerstein)

Abstract

We investigated how children's strategy selection on different problem types was influenced by whether two or three strategies were available in a computational estimation task. In addition, we examined the influence of working memory updating on these effects. Third and fourth graders (*N* = 725) were asked to indicate the best strategy for two-digit addition problems (e.g., 47 + 24) without calculating estimates. Homogeneous problems (i.e., unit digits both smaller or both larger than 5) and heterogeneous problems (i.e., one operand's unit digit smaller and the other larger than 5) were included. Children completed selection tasks under two conditions. First, a three-strategy condition, in which they could choose between the rounding-down (i.e., rounding both operands down), the mixed-rounding (i.e., rounding one operand down and the other up), and the rounding-up strategy (i.e., rounding both operands up). Second, a two-strategy condition, in which they could select between the rounding-down and the rounding-up strategy only.

Results revealed that children chose the best available strategy more often under the three-strategy than the two-strategy condition and chose the best strategy more often on homogeneous than heterogeneous problems. In addition, these effects were moderated by children's updating capacities. That is, children with less efficient updating showed worse selection performance on heterogeneous than on homogeneous problems under both conditions. In turn, children with more efficient updating obtained comparable performance for both problem types under both conditions. These findings have important implications to further our understanding of underlying processes in children's strategy selection in computational estimation.

Keywords: arithmetic; strategies; working-memory updating; computational estimation; elementary-school children.

Introduction

Computational estimation is an important component of mathematical cognition and provides information about the understanding of other mathematical concepts and the interplay between different mathematical procedures (e.g., Bisanz & LeFevre, 2013; Siegler & Booth, 2005; Star, Rittle-Johnson, Lynch, & Perova, 2009). Computational estimation can be defined as "the process of simplifying an arithmetic problem using some set of rules or procedures to produce an approximate but satisfactory answer through mental calculation" (LeFevre, Greenham, & Waheed, 1993, p. 95). These sets of procedures or strategies strive to achieve the principles of simplicity and proximity (LeFevre et al., 1993). That is, strategies should be used to simplify the arithmetic problem while the estimate should be as close to the exact solution as possible. In this regard, Xu, Wells, LeFevre, and Imbo (2014) have introduced the term of the optimal problem-based strategy which integrates the principles of simplicity and proximity; that is, rounding operands with unit digits smaller than five down to the nearest decade (e.g., 42 to 40) and rounding operands with unit digits larger than five up (e.g., 48 to 50). Following this logic, three different optimal problem-based rounding strategies would result for arithmetic problems with two operands: the rounding-down strategy (i.e., rounding both operands down to the nearest decades) for problems with both operands' unit digits smaller than five, the rounding-up strategy (i.e., rounding both operands up to the nearest decades) for problems with both operands' unit digits larger than five, and the mixed-rounding strategy (i.e., rounding one operand down and the other up to the nearest decades) for problems with one operand's unit digit smaller and the other larger than five.

Still, in a large number of previous studies on strategy use in computational estimation (e.g., Hammerstein, Poloczek, Lösche, Lemaire, & Büttner, 2019; Lemaire & Lecacheur, 2011), participants were constrained in their strategy use by only allowing the use of the rounding-down and the rounding-up strategy. The mixed-rounding strategy was excluded to increase difficulty of strategy selection and avoid ceiling effects (e.g., Lemaire & Lecacheur, 2011). Preventing participants from using the mixed-rounding strategy leads to a discrepancy between the instruction to use the *best* strategy (i.e., yielding the closest estimate) and the set of available strategies. That is, for so-called heterogeneous problems (i.e., unit digit of one operand smaller and the other larger than five) the optimal problem-based strategy (i.e., mixed-rounding) was not permitted, forcing participants to come up with "some alternative way of deciding which of the two available but non-optimal procedures would produce the answer closest to the exact answer" (Xu et al., 2014, p. 1483). This is reflected in strategy selection differences for homogeneous problems (i.e., unit digits of the operands either both smaller than five or both larger than five) and heterogeneous problems in that participants selected the best of the available strategies (yielding the closest estimate) more often on homogeneous than heterogeneous problems (e.g., Hammerstein et al., 2019; Lemaire & Lecacheur, 2011). Therefore, it can be concluded that the definition of the best strategy varies not only as a function of problem characteristics (i.e., unit digits) but also of task demands (i.e., which strategies are permitted). This probably influences participants' selection mechanisms.

Despite this evidence on the influence of available strategies in a task on participant's strategy choices, no study to date has directly compared selection performance when the optimal problem-based strategy is experimentally excluded for some problems to selection performance when all three strategies are allowed. Investigating the role of whether two or three strategies are available in a task for participants' strategy selection is crucial to further our understanding of underlying processes of selection mechanisms. The question at hand is, whether problem type differences (i.e., better strategy selection on homogeneous than heterogeneous problems) would also occur under a three-strategy condition or disappear in the presence of mixed-rounding being allowed. The absence of problem type differences when three strategies are permitted would strengthen the case of an optimal problem-based approach to strategy selection tasks. Consequently, it would point to the fact that restricting strategies may not only increase task difficulty but involves qualitatively different selection mechanisms that are not inherent to strategy selection in a more naturalistic setting such as in educational contexts.

Therefore, the goal of the present study was to determine whether children's strategy selection and problem type effects were influenced by whether they could choose between two or three available strategies. We asked third and fourth graders to indicate the best available strategy for two-digit addition problems (e.g., 47 + 24) without calculating the estimates. First, children worked on problems under a *three-strategy* condition, in which they could choose between the rounding-down, the mixed-rounding, and the rounding-up strategy. Then, they worked on problems under a *two-strategy* condition, in which they could choose only between the rounding-down and the rounding-up strategy. Half of the problems were homogeneous problems; the other half were heterogeneous problems. In accordance with the optimal problem-based approach, we expected selection performance to be facilitated when three strategies were permitted and problem type differences to occur only under the two-strategy condition.

Additionally, we assessed children's working memory updating with four different tasks. Research found updating to be a strong predictor of arithmetic performance (for an overview, see Bull & Lee, 2014) and strategy selection (Hammerstein et al., 2019). The present strategy selection tasks required children to monitor several pieces of information that are relevant for selecting the best available strategy. That is, crucial problem characteristics for strategy selection need to be identified (e.g., unit digits), rounding strategies must be activated in working memory and the best available strategy must be selected as a function of problem characteristics and task demands (e.g., different sets of strategies are permitted). As updating consists in "monitoring and coding incoming information for relevance to the task at hand" (Miyake et al., 2000, p. 57), we expected updating to facilitate the selection of both well-known optimal problem-based strategies and second-best strategies when the optimal problem-based strategy is not permitted.

Method

Participants

A total of 725 children were tested: The sample consisted of 449 third graders (231 males; age in months: M = 112.5, SD = 5.7; range = 91-133) and 276 fourth graders (149 males; age in months: M = 124.5, SD = 5.9; range = 113-153). The study was conducted near the end of the school year. Children were recruited from 50 different classes in 16 elementary schools in urban and suburban areas in the state of Hesse (Germany). The study was approved by the local ethics committee. Parents provided written informed consent and participants gave their verbal consent.

Material and Procedure

All participants completed the strategy selection tasks and four updating tasks (see supplementary material for detailed task descriptions). Children took part in three sessions, each lasting approximately 45 minutes, with an average of 12 days (SD = 11) between each session. Children were tested in groups of up to 11 individuals (*M* = 5, *SD* = 3). They solved all tasks on tablet computers. Tasks were programmed with .NET Framework 4.0. Instructions were presented verbally over headphones. Experimenters were present for each group to answer children's questions and provide further explanations if needed.

Strategy selection task. Children were presented with two-digit addition problems (e.g., 47 + 24; see supplementary material for detailed problem descriptions). They were asked to indicate the best available rounding strategy without calculating the estimates. First, children worked on problems under a three-strategy condition: they could choose between the rounding-down, the mixed-rounding, and the rounding-up strategy. On a second day, children worked on another set of problems under a two-strategy condition, for which they could choose only between the rounding-down and the rounding-up strategy. The three-strategy condition was always presented first to avoid carry-over effects from inhibiting the mixed-rounding strategy in the two-strategy condition on children's strategy selection in the three-strategy condition (Siegler & Lemaire, 1997).

Children were visually presented with the addition problems on the screen. Distinctive icons representing each of the available rounding strategies were displayed next to each problem: two arrows pointing down for the rounding-down strategy, one arrow pointing down and the other pointing up separated by a diagonal slash from an arrow pointing up and an arrow pointing down for the mixed-rounding strategy, and two arrows pointing up for the rounding-up strategy.

Under each condition, children worked on a set of 24 two-digit addition problems. To investigate problem type effects, half the problems were homogeneous problems and the other half were heterogeneous (i.e., mixed-unit) problems. No problem from one condition was the same as in the other condition. Still, problems from the two conditions were carefully matched regarding their combination of unit and decade digits to avoid systematic effects of different problem sets on children's performance.

In accordance with prior studies (e.g., Hammerstein et al., 2019; Lemaire & Lecacheur, 2011; Xu et al., 2014), the definition of the best available strategy varies as a function of available strategies. That is, under three-strategy condition, the best strategy corresponds to the optimal problem-based strategy for all problems: rounding-down for small-unit problems, mixed-rounding for mixed-unit problems, and rounding-up for large-unit problems (i.e., rounding down operands with unit digits smaller than five and rounding up operands with unit digits larger than five). Similarly, in the two-strategy condition, the best strategy was rounding-down for small-unit and rounding-up for large-unit problems. As the mixed-rounding strategy was not available under the two-strategy condition, rounding-down was the best available strategy for mixed-unit problems with sums of unit digits smaller than 10 and rounding-up for mixed-unit problems with sums of unit digits larger than 10 as that would yield the closer estimate.

Working memory updating tasks. Working memory updating was assessed with four different tasks to obtain a stable and purer measure of updating. These four tasks were the spatial keep-track, day keep-track, frog-position updating, and color updating tasks. Assessment of the tasks was distributed over the three test days. In the *spatial keep-track* task (Dirk & Schmiedek, 2016), children were presented with differently colored, fictitious creatures (i.e., monsters) at different positions on a 4 by 4 grid. Creatures disappeared and colored arrows were presented indicating covert movements of the corresponding creatures. Participants were asked to reproduce the final positions. In a *day keep-track* task (Hammerstein et al., 2019), participants were presented with either one, two, or three distinct calendar pages. On each page, the name for a weekday was written. Weekday names disappeared and a series of updating instructions (i.e., *1 forward*, *2 forward*, *1 backward* or *2 backward*, and corresponding arrows) was presented. Participants were asked to remember the updated weekdays for each calendar and to reproduce the final weekdays. In the *frog-position updating* task (LeFevre et al., 2009), a frog moved across distinct positions on an array of water lily leaves. Children were asked to indicate the last two to four positions of the frog. In the *color updating* task (Lee, Ng, Bull, Pe, & Ho, 2011), children were visually and verbally presented with sequences of colors. Children were asked to remember the last two or three colors. The percentages of correctly remembered items in each of the four tasks was used as a measure of individuals' updating. To obtain a task-nonspecific updating value, the factor scores of a factor analysis with a single factor were used for further analyses. We centered individuals' updating scores within grades; thus, a score of zero reflects the updating efficiency of an average child within his or her grade.

Results

Children's best strategy selection was examined with multilevel models (see supplementary material for a detailed description of the statistical method). Children were more likely to select the best available strategy under the three-strategy than under the two-strategy condition (PP = 81% vs. PP = 76%; β = 0.31, 95% CI [0.27, 0.34], $p < .001$) and were more likely to select the best strategy on homogeneous than on heterogeneous problems (PP = 85% vs. PP = 72%; β = -0.27, 95% CI [-0.30, -0.24], $p < .001$). The interaction of problem type and task condition (β = 0.52, 95% CI [0.49, 0.58], $p < .001$) was significant. In contrast to our expectation, this occurred because the problem type was even more important under the three-strategy than under the two-strategy condition. Moreover, children with more efficient updating were more likely to select the best strategy than children with less efficient updating (average predicted probability for children with a standardized updating score of +1 *SD*, PP_{+1SD} = 89% vs. PP_{-1SD} = 71%; β = 0.65, 95% CI [0.58, 0.73], $p < .001$) and fourth graders outperformed third graders (PP = 82% vs. PP = 76%; β = 0.25, 95% CI [0.18, 0.33], $p < .001$). Contrary to our expectation, the interaction of updating and task condition (β = 0.21, 95% CI [0.18, 0.24], $p < .001$) was significant as the number of available strategies was more important for children with more efficient than less efficient updating. As expected, the interaction of updating and problem type (β = -0.09, 95% CI [-0.12, -0.06], $p < .001$) was significant as children with more efficient updating outperformed children with less efficient updating specifically on heterogeneous problems. Importantly, effects were qualified by the three-way interaction of updating, task condition and problem type (β = 0.10, 95% CI [0.07, 0.13], $p < .001$). This occurred, because children with less efficient updating clearly obtained worse performance on heterogeneous than on homogeneous problems (58% vs. 84%), whereas children with more efficient updating were slightly more accurate on heterogeneous than homogeneous problems (91% vs. 86%). This not only occurred under the two-strategy but also under the three-strategy condition for which the problem type effect in children with less efficient updating was even larger (see Figure 1).

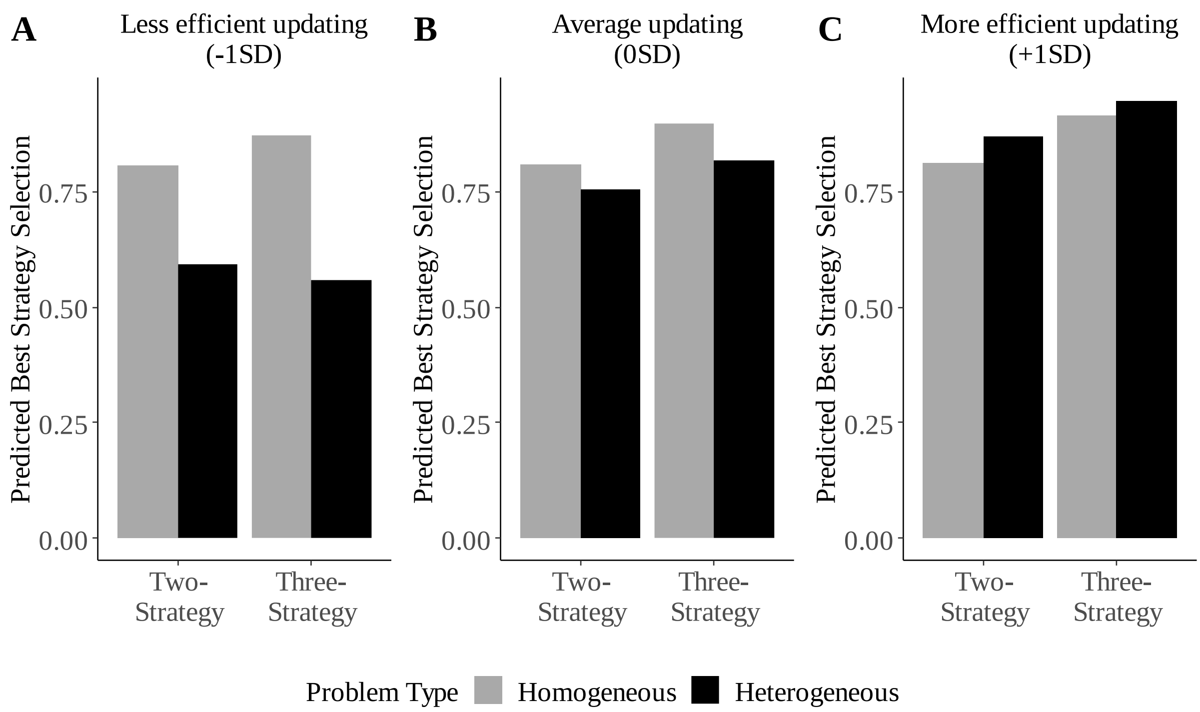


Figure 1. Influence of task condition and problem type on predicted best strategy selections of children with less efficient updating (A), average updating (B), and more efficient updating (C).

Discussion

The present study investigated the influence of the number of available strategies and working memory updating on third and fourth graders' strategy selection and moderating influences on problem type effects. Large samples of third and fourth graders were asked to indicate the best available rounding strategy for two-digit addition problems without executing the strategy. The results show clearly that updating and number of available strategies have relevant effects on strategy choice in computational estimation.

Numerous studies in computational estimation have restricted participants in their strategy selection (e.g., Hammerstein et al., 2019; Lemaire & Lecacheur, 2011). However, if participants base their strategy selection on each operand's unit digit separately (i.e., the optimal problem-based procedure; Xu et al., 2014), selection on heterogeneous problems with a restricted strategy set requires cognitive processes that are not inherent to strategy selection in more naturalistic settings. In line with this hypothesis, the present data showed that children were more likely to select the best available strategy when they could choose between three strategies than when they could choose between two strategies. This indicates that strategy selection was more accurate when the optimal problem-based strategy was available for all problems rather than only a subset of problems.

Interestingly and uniquely, we found that relative strategy selection performance as a function of problem type and available strategies differed with children's efficiency of updating. That is, children with more efficient updating performed equally well on heterogeneous and homogeneous problems in both conditions. It is not surprising, that children with more efficient updating show similar selection performance for both problem types when three strategies are available as they have likely established stable associations between problem characteristics and the optimal problem-based strategy (see also Hammerstein et al., 2019). In addition, these children selected the best available strategy equally often on homogeneous and heterogeneous when only two strategies were permitted. This finding is different to problem type effects in previous research with a restricted strategy set, which reported better strategy selection for homogeneous than heterogeneous problems specifically for older children and children with more efficient updating (Hammerstein et al., 2019; Lemaire & Lecacheur, 2011)[4]. We argue that this difference likely occurred due to changes to task demands. In the present study, children were asked to only select the best strategy without executing it and the task-instruction solely focused on accuracy of selection rather than speed. Focusing the task-goal on accuracy of strategy use leads adults to obtain better strategy selection performance when three strategies are permitted (Xu et al., 2014). Even though adults perform equally well when the focus is on just accuracy or accuracy and speed, the clear focus on accuracy might be even more important in children and when only two strategies are permitted. This highlights the importance of updating for children to figure out the better strategy even for problems for which the optimal problem-based strategy was excluded. Additionally, the present results imply that this selection involves more complex processes. Children with more efficient updating specifically took more time to identify the best available strategy for heterogenous problems in comparison to homogeneous problems under two-strategy condition (+1.2 s; see supplementary material for the analysis of children's response times). This

[4] At the same time, findings on age-related improvements and main effects of working memory updating could be replicated (e.g., Hammerstein et al., 2019; Lemaire & Lecacheur, 2011) and extended to strategy selection performance when only selection is required.

indicates that these children fulfilled the demands of the more complex processes to identify the second-best strategy by more deeply analyzing crucial problem features.

Another interesting result is that children with less efficient updating clearly obtained worse performance on heterogeneous than on homogeneous problems both under two-strategy and three-strategy conditions. This means that these children were not only less able to accomplish the demanding processes of selecting the second-best strategy on heterogeneous problems when mixed-rounding was excluded but also to identify mixed-rounding as optimal problem-based strategy for heterogeneous problems when three strategies were available. An explanation might be that these children had not yet created a stable association between heterogeneous problems and the mixed-rounding strategy. Participants' working memory capacity, which is closely linked to updating (e.g., Ecker, Lewandowsky, Oberauer, & Chee, 2010), contributes to establishing associations in long-term memory and to accessing existing associations (Unsworth, 2016). Therefore, children with less efficient updating might have had difficulties in forming or retrieving associations for heterogeneous problems regardless of the number of available strategies. Another explanation might be that these children had difficulties in considering unit digits of both operands simultaneously and based their strategy selection mainly on the first or the second unit digit to save cognitive demands of holding both unit digits in memory during the selection process.

In sum, results showed that efficient updating processes enabled children to select the best available strategy specifically on more difficult, heterogeneous problems under both task conditions. This indicates that updating enabled children to (a) accomplish the processes required to select the second-best strategy in a restricted set when the task-goal is to be accurate and (b) consider operands' unit digits jointly even when they require different rounding procedures. In turn, this might imply that children with less efficient updating likely have established associations of the optimal problem-based strategy as proposed by Xu et al. (2014) only for single digits but not for the combination of different procedures in problems with multiple operands. Therefore, in educational contexts, it might be recommendable to adjust instructions or materials to account for the difficulties of these children. As proposed by other authors (Jitendra et al., 2007; Lemaire, Luwel, & Brun, 2017), different strategies should be instructed and practiced one at a time in children with fewer cognitive resources. However, later lessons should also focus on the simultaneous consideration of multiple stimuli, as this seems to be specifically difficult for these children.

The present findings emphasize the role of strategy sets and problem characteristics for strategy selection mechanisms. Researchers in the field of strategies should be aware that alterations of these variables might not solely influence task difficulty but likely require different cognitive processes that are not inherent to strategy selection in more naturalistic settings. This might even shed new light onto findings of previous studies with a restricted strategy set and without instructional focus strictly on accuracy (e.g., Hammerstein et al., 2019; Lemaire & Lecacheur, 2011). That is, the influence of individual differences might have been underestimated in terms that the capability of children with more efficient updating to identify the better of the available strategies was even higher than these studies found. Finally, existing models of strategy selection (Payne et al., 1993; Rieskamp & Otto, 2006; Shrager & Siegler, 1998) could be expanded to make assumptions on selection processes when a restricted strategy repertoire is available. Future studies with verbal protocols could be conducted to get a more detailed picture of the specific procedures underlying children's strategy selection when different strategy sets are available and different problem types are presented and when the focus of task-instruction varies.

References

Bisanz, J., & LeFevre, J.-A. (2013). Strategic and nonstrategic processing in the development of mathematical cognition. In D. F. Bjorklund (Ed.), *Children's Strategies: Contemporary Views of Cognitive Development* (pp. 213–243). Psychology Press.

Bull, R., & Lee, K. (2014). Executive functioning and mathematics achievement. *Child Development Perspectives*, *8*(1), 36–41. https://doi.org/10.1111/cdep.12059

Dirk, J., & Schmiedek, F. (2016). Fluctuations in elementary school children's working memory performance in the school context. *Journal of Educational Psychology*, *108*(5), 722–739. https://doi.org/10.1037/edu0000076

Ecker, U. K. H., Lewandowsky, S., Oberauer, K., & Chee, A. E. H. (2010). The components of working memory updating: An experimental decomposition and individual differences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*(1), 170–189. https://doi.org/10.1037/a0017891

Hammerstein, S., Poloczek, S., Lösche, P., Lemaire, P., & Büttner, G. (2019). Effects of working memory updating on children's arithmetic performance and strategy use: A study in computational estimation. *Journal of Experimental Child Psychology*, *184*, 174–191. https://doi.org/10.1016/j.jecp.2019.04.003

Jitendra, A. K., Griffin, C. C., Haria, P., Leh, J., Adams, A., & Kaduvettoor, A. (2007). A comparison of single and multiple strategy instruction on third-grade students' mathematical problem solving. *Journal of Educational Psychology*, *99*(1), 115–127. https://doi.org/10.1037/0022-0663.99.1.115

Lee, K., Ng, S. F., Bull, R., Pe, M. L., & Ho, R. H. M. (2011). Are patterns important? An investigation of the relationships between proficiencies in patterns, computation, executive functioning, and algebraic word problems. *Journal of Educational Psychology*, *103*(2), 269–281. https://doi.org/10.1037/a0023068

LeFevre, J.-A., Greenham, S. L., & Waheed, N. (1993). The Development of Procedural and Conceptual Knowledge in Computational Estimation. *Cognition and Instruction*, *11*(2), 95–132. https://doi.org/10.1207/s1532690xci1102_1

LeFevre, J.-A., Skwarchuk, S.-L., Smith-Chant, B. L., Fast, L., Kamawar, D., & Bisanz, J. (2009). Home numeracy experiences and children’s math performance in the early school years. *Canadian Journal of Behavioural Science*, *41*(2), 55–66. https://doi.org/10.1037/a0014532

Lemaire, P., & Lecacheur, M. (2011). Age-related changes in children's executive functions and strategy selection: A study in computational estimation. *Cognitive Development.* https://doi.org/10.1016/j.cogdev.2011.01.002

Lemaire, P., Luwel, K., & Brun, F. (2017). Does the Number of Available Strategies Change How Children Perform Cognitive Tasks? Insights from Arithmetic. *Journal of Educational and Developmental Psychology*, *7*(2), 43. https://doi.org/10.5539/jedp.v7n2p43

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

Payne, J. W., Johnson, E. J., & Bettman, J. R. (1993). *The adaptive decision maker*. Cambridge University Press. https://doi.org/10.1017/CBO9781139173933

Rieskamp, J., & Otto, P. E. (2006). Ssl: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, *135*(2), 207–236. https://doi.org/10.1037/0096-3445.135.2.207

Shrager, J., & Siegler, R. S. (1998). SCADS: A Model of Children's Strategy Choices and Strategy Discoveries. *Psychological Science*, *9*(5), 405–410. https://doi.org/10.1111/1467-9280.00076

Siegler, R. S., & Booth, J. L. (2005). Development of numerical estimation. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 197–212). Psychology Press.

Siegler, R. S., & Lemaire, P. (1997). Older and younger adults' strategy choices in multiplication: Testing predictions of ASCM using the choice/no-choice method. *Journal of Experimental Psychology: General*, *126*(1), 71–92. https://doi.org/10.1037/0096-3445.126.1.71

Star, J. R., Rittle-Johnson, B., Lynch, K., & Perova, N. (2009). The role of prior knowledge in the development of strategy flexibility: the case of computational estimation. *ZDM*, *41*(5), 569–579. https://doi.org/10.1007/s11858-009-0181-9

Unsworth, N. (2016). Working memory capacity and recall from long-term memory: Examining the influences of encoding strategies, study time allocation, search efficiency, and monitoring abilities. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *42*(1), 50–61. https://doi.org/10.1037/xlm0000148

Xu, C., Wells, E., LeFevre, J.-A., & Imbo, I. (2014). Strategic flexibility in computational estimation for Chinese- and Canadian-educated adults. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *40*(5), 1481–1497. https://doi.org/10.1037/a0037346

Supplementary Material

A Detailed Description of the Strategy Selection Tasks

Tasks consisted of a set of 24 two-digit addition problems for the *three-strategy* condition and a set of 24 two-digit addition problems for the *two-strategy* condition. No problem from one condition was the same as in the other condition. Still, problems from the two conditions were carefully matched regarding their combination of unit and decade digits to avoid systematic effects of different problem sets on children's performance. Problems included operands ranging from 17 to 94; exact sums ranged from 46 to 166. Each set consisted of 12 so-called homogeneous problems (i.e., problems for which unit digits of operands are either both smaller or both larger than five) and 12 so-called heterogeneous or mixed-unit problems (i.e., problems for which the unit digit was smaller than five in one operand and larger than five in the other operand). Half of the homogeneous problems were small-unit problems (i.e., problems for which unit digits of both operands were smaller than five, such as 32 + 21) and the other half large-unit problems (i.e., problems for which unit digits of both operands were larger than five, such as 37 + 19). For half of the mixed-unit problems, the unit digit was smaller than five in the first operand and larger than five in the second operand (e.g., 53 + 28) and the reverse in the other problems (e.g., 47 + 24). The first addend was larger than the second addend in half the problems of each problem type and smaller in the other problems. Additionally, within each of these halves, the unit digit of the first addend was larger than the unit digit of the second addend in half the problems and the reverse in the other problems. Moreover, half of the problems of each problem type required a carry over 100 (e.g., 38 + 94). For each participant, problems were presented in random order. Finally, based on previous findings in arithmetic (see Cohen Kadosh & Dowker, 2015; Gilmore et al., 2018, for overviews), we controlled the following factors: (1) no operands had a 0 or a 5 as unit digit (e.g., 20 + 63; 25 + 63), (2) unit and decade digits were never the same within operands (e.g., 44 + 23), (3) unit or decade digits were never the same for both operands in a given problem (e.g., 32 + 62; 49 + 41), (4) no problems with reverse order of the operands from another problem were presented (e.g., if 68 + 24 was included, 24 + 68 was not), and (5) no operand had its closest decade equal to 0, 10, or 100.

B Detailed Description of the Working Memory Updating Tasks

Working-memory updating was assessed with four different tasks to obtain a stable and purer measure of updating. These four tasks were the spatial keep-track, day keep-track, frog-position updating, and color updating tasks. In the *spatial keep-track* task (Dirk & Schmiedek, 2016), children were presented with either two or three differently colored, fictitious creatures (i.e., monsters) at different positions on a 4 by 4 grid. Children were asked to remember the initial positions of the creatures. The creatures disappeared after 3000 ms. Then, arrows matching the creatures' colors were successively presented in the center of the grid for 2500 ms with inter-stimulus-intervals of 500 ms. Each arrow indicated that the creature of the corresponding color should covertly move one position in the arrow's direction. Three arrows for two-creature trials and four arrows for three-creature trials were presented (i.e., three and four updating operations, respectively). Participants were asked to remember the new, updated positions of all creatures and to reproduce the final positions on an empty grid at the end of each trial. Each of the two difficulty levels (i.e., two vs. three creatures) was presented eight times.

An arithmetic keep-track task (Dirk & Schmiedek, 2016) was adapted and changed into a *day keep-track* task (Hammerstein et al., 2019) to avoid an arithmetic confound in our measures of updating. Participants were presented with either one, two, or three distinct calendar pages. On each page, the name for a weekday was written. Children were asked to remember the initial weekdays on the calendar pages. The names of the weekdays appeared for 3000 ms for one-calendar trials and for 4000 ms for two- and three-calendar trials. After displaying weekday names, a series of updating instructions (i.e., German words for *1 forward*, *2 forward*, *1 backward* or *2 backward,* and corresponding arrows) was presented on empty calendar pages for 3000 ms, one at a time. Updating operations were presented with inter-stimulus-intervals of 500 ms. Participants were asked to remember the updated weekdays for each calendar and to reproduce the final weekdays. The weekdays had to be updated three times for one- and two-calendar and three or four times for three-calendar trials. Each of the three difficulty levels (i.e., one, two, or three-calendar trials) was presented six times resulting in a total of 18 trials.

Next, in the *frog-position updating* task (LeFevre et al., 2009), a frog moved across 3-7 distinct positions on an array of eight irregularly positioned water lily leaves. Each position of the frog was displayed for 1750 ms with inter-stimulus-intervals of 250 ms. Then, children were asked to indicate either the last two, three, or four positions of the frog. The task required updating as children were not told after how many positions the sequence would stop. They worked on five trials with the last two positions, five trials with the last three positions, and five trials with the last four positions to remember.

Finally, children worked on 16 trials of a *color updating* task (Lee, Ng, Bull, Pe, & Ho, 2011). They consisted of sequences of 4-8 colors, one color at a time. Each color was presented visually for 1500 ms, simultaneously the color was labelled by a female voice. The color card was covered for 500 ms and followed by the presentation of the next color. Children were asked to always keep either the last two or three colors in memory and indicate them after each trial. Again, children were not told after how many colors the sequence would stop. They worked on eight trials with the last two colors and eight trials with the last three colors to remember. Stimuli consisted of eight colors with monosyllabic names.

The percentages of correctly remembered items in each of the four tasks was used as a measure of individuals' efficiency of working memory updating. If more than half of the trials of a difficulty level were missing for a participant due to task abortions or technical errors, the subtest score was not calculated. Thus, subtest scores were calculated for 721, 720, 722 and 711 children out of 725 children for the spatial keep-track, the day keep-track, the frog updating, and the color updating tasks, respectively. The internal consistency of the updating tasks, as calculated by Cronbach's alpha for four balanced parcels each was acceptable to high (i.e., α = .93, .81, .88, .80 for the spatial keep-track, day keep-track, frog updating, color updating tasks, respectively).

A common problem in investigating the influence of EF is the so-called task-impurity problem (Rabbitt, 1997). That is, executive tasks contain task-specific variance due to non-executive abilities such as verbal or visuo-spatial abilities. We addressed this problem by extracting factor scores of a latent factor underlying our four updating tasks. With this approach we obtained an updating-specific factor and minimized the influence of task-specific processes within each task. We centered individual's updating scores within grades to investigate the influence of updating beyond age-related effects; thus, a score of zero reflects the efficiency of updating of an average child within his or her grade.

C Statistical Method for the Analysis of Children's Best Strategy Selection

Children's best strategy selection was examined with multilevel models. The data were modeled with a two-level structure[5], with trials on Level 1 nested within 725 participants on Level 2. Children's dichotomous strategy selection variable was analyzed with logit models for binary responses. To be able to interpret effects in terms of probabilities that y = 1 (e.g., that children selected the best strategy on a given problem), we calculated the predicted probabilities of children's strategy selection for the different values of predictors and report them in the following as percentages. We calculated the predicted probabilities as population-averaged probabilities by averaging over different simulated Level 2 random intercepts. Thus, percentages can be interpreted as predicted probabilities that are true for the average of children with a certain predictor combination (see also Steele, 2009, for the calculation of predicted probabilities in logit models).

D Analysis of Children's Response Times

Introduction. Children's response times were analyzed as an additional parameter to assess difficulty of selection procedures[6]. That is, we expected children to take longer for identifying the best available strategy on problems with more complex procedures, such as the best available strategy for heterogeneous problems when mixed-rounding is excluded.

Method. Children's response times to select strategies were examined with multilevel models (see Table D1 for the complete depiction of statistical effects). The data were modelled with a two-level structure, with trials on Level 1 nested within 725 participants on Level 2. Only reaction times for trials on which children selected the best strategy were included in the

[5] Including schools and classes as further levels did not change the original results and analyses revealed clearly smaller variances in random effects that were accounted for by schools (5% of the total random effect variance) or by classes within schools (2% of the total random effect variance) than they were accounted for by individuals (47% of the total random effect variance). Additionally, we did not measure any predictors on school or class level but only on individual level. Therefore, we omitted these levels from the multilevel analysis.
[6] Note that we did not include this analysis in the main body of the manuscript because it would have exceeded the length of a brief report.

analysis. To omit trials with extreme individual response times, *z*-standardized values of response times were calculated within participants. Trials were excluded from estimation time analyses if the values exceeded 2.5 because they most likely resulted from children's transient distraction from the task.

Results. Children with more efficient updating were faster than children with less efficient updating (2.9 s vs. 3.4 s; $p < .001$). Children were faster when two strategies were available than when three strategies were available (3.0 s vs. 3.3 s; $p < .001$) and were faster on homogeneous than heterogeneous problems (2.9 s vs. 3.4 s; $p < .001$). The interaction between task condition and problem type was significant ($p < .001$). This occurred because the heterogeneous-homogeneous problem type difference was larger under the two-strategy (3.4 s vs. 2.6 s) than the three-strategy condition (3.3 s vs. 3.2 s). This effect was even larger in children with more efficient updating, shown by the three-way interaction between updating, problem type and task condition was significant ($p < .001$). That is, specifically children with more efficient updating took longer to select the best strategy for heterogeneous problems under the two-strategy condition. Under the two-strategy condition these children took on average 1.2 s longer on heterogeneous than homogeneous problems whereas this difference was only 0.1 s under three-strategy condition. For children with less efficient updating the heterogeneous-homogeneous difference was 0.4 s with two available strategies and no difference occurred with three strategies.

Discussion. These findings imply that children took longest to identify the second-best strategy for heterogeneous problems when the mixed-rounding strategy was not allowed. This was specifically true for children with more efficient updating. These results indicate that indeed selection of the second-best strategy involves more complex procedures and that children with more efficient updating took more time on these problems to follow closely the instruction to be accurate.

Table D1

Statistical effects of a multilevel model examining children's response times

Fixed Part	**β**	**95% CI**	***p***
Intercept	3.13	[3.00, 3.27]	< .001
Grade	-0.03	[-0.16, 0.11]	.71
Working memory updating	-0.24	[-0.37, -0.11]	< .001
Task condition	0.12	[0.09, 0.15]	< .001
Problem type	0.22	[0.19, 0.24]	< .001
Task Condition x Problem Type	-0.19	[-0.22, -0.16]	< .001
Working memory updating x Task Condition	-0.17	[-0.20, -0.14]	< .001
Working memory updating x Problem Type	0.09	[0.06, 0.12]	< .001
Working memory updating x Task Condition x Problem Type	-0.07	[-0.10, -0.04]	< .001
Random Part	**Variance**		
Participant	3.0		
Residual	5.6		

References

Cohen Kadosh, R., & Dowker, A. (2015). *Oxford Handbook of Numerical Cognition*. Oxford: Oxford University Press.

Dirk, J., & Schmiedek, F. (2016). Fluctuations in elementary school children's working memory performance in the school context. *Journal of Educational Psychology*, *108*(5), 722–739. https://doi.org/10.1037/edu0000076

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

Hammerstein, S., Poloczek, S., Lösche, P., Lemaire, P., & Büttner, G. (2019). Effects of working memory updating on children's arithmetic performance and strategy use: A study in computational estimation. *Journal of Experimental Child Psychology*, *184*, 174–191. https://doi.org/10.1016/j.jecp.2019.04.003

Lee, K., Ng, S. F., Bull, R., Pe, M. L., & Ho, R. H. M. (2011). Are patterns important? An investigation of the relationships between proficiencies in patterns, computation, executive functioning, and algebraic word problems. *Journal of Educational Psychology*, *103*(2), 269–281. https://doi.org/10.1037/a0023068

LeFevre, J.-A., Skwarchuk, S.-L., Smith-Chant, B. L., Fast, L., Kamawar, D., & Bisanz, J. (2009). Home numeracy experiences and children's math performance in the early school years. *Canadian Journal of Behavioural Science*, *41*(2), 55–66. https://doi.org/10.1037/a0014532

Rabbitt, P. (Ed.). (1997). *Methodology of frontal and executive function*. Hove: Psychology Press.

Steele, F. (2009). *Module 7: Multilevel models for binary responses. Concepts*. LEMMA VLE, Centre for Multilevel Modelling. Retrieved from http://www.cmm.bris.ac.uk/lemma/course/view.php?id = 13

6.4 Study 4: Effects of Presentation Conditions on Strategy Use

Hammerstein, S., Poloczek, S., Lösche, P., Lemaire, P., & Büttner, G. (2020). Effects of presentation duration and modality on children's strategy selection in a computational estimation task. *Manuscript submitted for publication.*

Effects of working memory updating on children's arithmetic performance and strategy use: A study in computational estimation

Svenja Hammerstein[1,2]*, Sebastian Poloczek[1,2], Patrick Lösche[1,2], Patrick Lemaire[3], & Gerhard Büttner[1,2]

[1]Goethe Universität Frankfurt am Main, Germany

[2]Centre for Individual Development and Adaptive Education of Children at Risk (IDeA), Frankfurt am Main, Germany

[3]Aix-Marseille Université, LPC, & CNRS, Marseille, France

*Corresponding author at: Goethe Universität Frankfurt am Main, Theodor-W.-Adorno-Platz 6, 60629 Frankfurt am Main; Tel.: +49 69 798 35377; E-mail address: hammerstein@psych.uni-frankfurt.de (S. Hammerstein)

Abstract

How do presentation conditions influence children's strategy selection and execution? To address this issue, we ran two experiments in which third and fourth graders were asked to find estimates for two-digit addition problems (e.g., 52 + 39). Children were tested in three presentation groups: (1) time-unlimited visual, (2) time-limited visual, or (3) time-limited auditory conditions. Moreover, we assessed children's working memory updating and arithmetic fluency. In Experiment 1, children could choose between the rounding-down, rounding-up, or mixed-rounding strategies to estimate each problem. In Experiment 2, each child was cued to use the best strategy on each problem. Regarding strategy selection (Expt. 1), children used the available strategies more flexibly when problems were visually displayed with no time limitation as well as when problems were auditorily presented. Also, children were most likely to choose the best strategy when problems were presented visually with no time limitation. Regarding strategy execution (Expt. 2), children were fastest to estimate when problems were visually displayed with time limitation and executed strategies most accurately when problems were visually displayed with no time limitation. Moreover, effects of presentation conditions were moderated by updating. Best strategy selection increased most under time-unlimited presentation for children with more efficient updating capacities. Also, children with more efficient updating were equally fast under visual and auditory presentation. These findings have important implications for determining conditions under which children select the best strategy on each problem most often and execute strategies most efficiently, as well as for understanding mechanisms underlying strategic behaviour.

Keywords: arithmetic; strategies; working memory updating; presentation modality; presentation duration; computational estimation

Introduction

In the vast majority of studies on arithmetic, participants are presented with problems visually and with no time restriction (e.g., Hammerstein, Poloczek, Lösche, Lemaire, & Büttner, 2019; LeFevre, Greenham, & Waheed, 1993; Lemaire & Lecacheur, 2011). While this is a legitimate study design in experimental research, children are often confronted with cognitive tasks of various presentation formats, modalities, and durations in real-life. That is, problem display duration might be limited, or presentation of problems might be auditory rather than visual. A few previous studies found that arithmetic performance varies under different presentation conditions (Adams & Hitch, 1997; Fürst & Hitch, 2000; Klingner, Tversky, & Hanrahan, 2011; LeFevre, Lei, Smith-Chant, & Mullins, 2001). However, the mechanisms responsible for these variations remain largely unknown. Here, we adopt a strategy-focused approach to address this issue and aim at determining whether variations in presentation conditions are accompanied by systematic variations in how children accomplish cognitive tasks (i.e., differences in strategies) and whether effects of presentation conditions are moderated by individual differences in cognitive processes. The present study addresses these issues in two experiments by manipulating presentation duration and presentation modality while third- and fourth-grade children with different levels of working memory updating worked on computational estimation tasks. We review the literature on previous findings on the influence of presentation conditions on arithmetic performance before outlining the logic of our present study.

Influence of Presentation Conditions on Performance in Mental Arithmetic

LeFevre and colleagues (2001) asked young adults to solve single-digit multiplication problems. Problems were auditorily presented in one condition and visually displayed in another condition. Participants made more errors when problems were auditorily presented. Error patterns suggested that participants activated a phonological code under auditory presentation, which interfered with the arithmetic process. Similarly, Klingner and colleagues (2011) found more errors and task-evoked pupil dilations when undergraduates solved multiplication problems under auditory condition than under visual condition. The authors interpreted performance benefits under visual presentation as stemming from poststimulus visual persistence alleviating cognitive load from working memory. In both studies, presentation duration was matched for the two presentation conditions. Another study systematically varied the presentation duration of visually displayed tasks. Fürst and Hitch (2000) asked young adults to solve three-digit addition problems under a brief visual presentation condition (i.e., problems were displayed for 4000 ms) and under a time-unlimited, visual presentation condition (i.e., problems were displayed until participants provided their answers). Participants made more errors in the brief presentation condition. No difference in response times was found. The authors proposed that a longer duration enhances the maintenance of the problem in working memory and frees cognitive resources for the arithmetic process.

To our knowledge, there is currently only one published study on the influence of presentation conditions on arithmetic performance in children, Adams and Hitch (1997) presented 7- to 11-year-olds with multi-digit addition problems. Children obtained better performance (i.e., larger addition spans and faster response times) when problems were visually displayed with no time restriction as compared to when problems were auditorily presented. The authors proposed that better performance under visual presentation resulted from the visual display providing an “external record” (p. 32), which reduced working memory load. Note that the presentation duration was not matched for auditory and visual conditions.

Overall, previous studies found effects of presentation conditions (modality and/or duration) on participants’ arithmetic performance. Results are consistent with assuming that a (permanent) visual display activates representations of the task that are beneficial for the arithmetic process as they free cognitive resources. However, these studies were only designed to test differences in arithmetic performance and do not make assumptions on whether performance differences are accompanied by strategic variations. Therefore, little is known about the source of performance differences under different presentation conditions. Do participants use different strategies, or do they use the same set of strategies but execute them with different levels of efficacy? Results reported by Lemaire and Brun (2016) suggest that participants may use different mechanisms under different presentation condition, as they found that participants used the better strategy more often to estimate sums of addition problems under a long (~4 s) than under a short presentation condition (~2 s). Xu, Wells, LeFevre, and Imbo (2014) reported opposite findings but also found participants to use different strategies under different conditions. That is, a short presentation (2 s) encouraged undergraduates to select the best strategy more often than a long presentation (4 s and 6 s).

Previous works (a) showed that children use several strategies to accomplish computational estimation tasks (for overviews, see Lemaire, 2018; Siegler, 2007) and (b) documented that differences in strategy use are influenced by individual differences in cognitive capacities (Hammerstein et al., 2019; Lemaire & Lecacheur, 2011). However, it is currently unknown whether children would use different strategies and/or execute strategies with different levels of efficacy under varying presentation conditions while accomplishing a cognitive task, like the computational estimation task tested here. It is also unknown whether effects of presentation duration and modality interact with participants' cognitive resources (e.g., working memory updating), such that some conditions are more beneficial for some individuals and more detrimental for other individuals. The present study aims to address these issues. We pursued this goal in the domain of computational estimation because it is an important component of mathematical cognition and has been the subject of many previous experimental studies to investigate strategic behaviour (e.g., Bisanz & LeFevre, 2013; Dowker, 1997; Dowker, Flood, Griffiths, Harriss, & Hook, 1996; Lemaire, Arnaud, & Lecacheur, 2004; Siegler & Booth, 2005; Sowder & Wheeler, 1989; Star, Rittle-Johnson, Lynch, & Perova, 2009).

The Present Study

In two experiments, we asked third and fourth graders to find approximate sums for two-digit addition problems (e.g., 52 + 39). In Experiment 1, we investigated whether children would use different strategies under different presentation conditions. For this purpose, children could choose among three available strategies: the rounding-down strategy (i.e., rounding both operands down to the nearest decades, like doing 50 + 30 = 80 to estimate 52 + 39), the rounding-up strategy (i.e., rounding both operands up to the nearest decades, like doing 60 + 40 = 100), or the mixed-rounding strategy (i.e., rounding one operand down and the other up to the nearest decades, like doing 50 + 40 = 90). In Experiment 2, we investigated whether children's efficacy of strategy execution would differ with varying presentation condition when children are directed which strategy to use. For this purpose, the best strategy was cued on each problem under a no-choice condition (e.g., rounding down was cued on problems with both operand's unit digits smaller than 5, such as 52 + 34).

To investigate effects of presentation conditions, children of both experiments were randomly assigned to one of three groups: (1) a time-unlimited visual presentation condition, (2) a time-limited visual presentation condition, and (3) a time-limited auditory presentation condition. In the time-unlimited visual condition, problems were displayed on the screen until children provided their response. In the time-limited auditory condition, problems were auditorily presented via headphones. In the time-limited visual condition, problems were displayed for 3000 ms. This design enabled us to (a) compare the present findings with previous findings and control that in our present sample children approach the task like in previous studies, (b) examine effects of presentation conditions as varying along the presentation-time dimension (by comparing time-limited and time-unlimited visual presentation condition), and (c) test effects of presentation conditions as varying along the presentation-modality dimension (by contrasting visual and auditory time-limited conditions).

We hypothesized that an unlimited presentation would facilitate children in their strategy use as they have more time to deeply analyse crucial problem features (e.g., size of unit digits) and as Fürst and Hitch (2000) noted, this would lessen the burden on their working memory because external representations would be available throughout the task. In addition, we assumed that the activation of phonological codes due to auditory presentation likely interferes with strategy selection and execution processes (see also Campbell, 1994; LeFevre et al., 2001). Therefore, we expected children's best strategy selection (Expt. 1) and strategy execution (Expt. 2) to benefit from (a) time-unlimited relative to time-limited presentation and (b) visual relative to auditory presentation.

Prior works have suggested that different presentation conditions impose different cognitive demands on working memory. One working memory component, namely working memory updating, has been found to influence arithmetic performance (for an overview, see Bull & Lee, 2014), including strategic behaviour in a computational estimation task tested here (Hammerstein et al., 2019). In the present study, we assessed children's working memory updating with four tasks to test the hypothesis that more demanding presentation conditions (i.e., time-limited visual and auditory presentations) would negatively affect strategy selection

(Expt. 1) and strategy execution (Expt. 2) specifically of children with less efficient working memory updating.

General Method

In this study, we used the choice/no-choice method (Siegler & Lemaire, 1997) to investigate children's strategy selection and execution in computational estimation. In the choice condition (Expt. 1), children could choose between three different rounding strategies for each problem. In the no-choice condition (Expt. 2), for each problem children were cued with the strategy to use.

Experiment 1: Strategy Selection

Method

Participants. A total of 347 children were tested: The sample consisted of third graders (108 males; age in months: *M* = 112.1, *SD* = 4.9; range = 101-128) and 140 fourth graders (78 males; age in months: *M* = 123.8, *SD* = 5.0; range = 114-142). The study was conducted near the end of the school year. Children were recruited from 50 different classes in 16 primary schools in urban and suburban areas in the state of Hesse (Germany). The study was approved by the local ethics committee. Furthermore, parents provided written informed consent and participants gave their verbal consent.

Procedure. Participants completed the computational estimation task, an arithmetic fluency task, and four updating tasks. Children took part in three sessions, each lasting approximately 45 minutes, with an average of 12 days (*SD* = 11) in-between each session. Children were tested in groups of up to 11 individuals (*M* = 5, *SD* = 3). They solved all tasks on tablet computers. Tasks were programmed with .NET Framework 4.0. Instructions were presented verbally over headphones. Experimenters were present for each group to answer children's questions and provide further explanations if needed.

Material.

Computational estimation task.

Stimuli. Children worked on two blocks of 18 two-digit addition problems each. Problems included operands ranging from 21 to 89, with exact sums ranging from 58 to 163. Each block consisted of 6 small-unit problems (i.e., problems for which unit digits of both operands were smaller than 5, such as 71 + 32), 6 large-unit problems (i.e., problems for which unit digits of both operands were larger than 5, such as 49 + 27), and 6 mixed-unit problems (i.e., problems for which the unit digit was smaller than 5 in one operand and larger than 5 in the other operand, such as 34 + 49). The unit digit was smaller than 5 in the first operand and larger than 5 in the second operand (e.g., 34 + 49) in half the mixed-unit problems, and the reverse in the other mixed-unit problems (e.g., 38 + 21). The first summand was larger than the second summand in half the problems of each problem type and smaller in the other problems. Additionally, the unit digit of the first summand was larger than the unit digit of the second summand in half the problems and the reverse in the other problems. Moreover, half the problems of each problem type involved a carry over 100 (e.g., 86 + 42). For each participant, problems were presented in random order.

Based on previous findings in arithmetic (for overviews, see Cohen Kadosh & Dowker, 2015; Gilmore, Göbel, & Inglis, 2018), we controlled the following factors: (1) no operands had a 0 or a 5 as unit digit (e.g., 20 + 63; 25 + 63), (2) unit and decade digits were never the same within operands (e.g., 44 + 23), (3) unit or decade digits were never the same for both operands in a given problem (e.g., 32 + 62; 49 + 41), (4) no problems with reverse order of the operands from another problem were presented (e.g., if 68 + 24 was included, 24 + 68 was not), and (5) no operand had its closest decade equal to 0, 10, or 100.

Children were randomly assigned to one of three presentation conditions: (1) a time-unlimited visual group, (2) a time-limited visual group, or (3) a time-limited auditory group. In the time-unlimited visual group, children saw each problem until they provided their answer. In the time-limited auditory group, children were verbally presented with the addition problems via headphones. The audio presentations were recorded by a native German speaker. The spoken stimuli were adjusted by editing the wave files to last for 3000 ms as this was the mean time to read problems aloud in pilot work. In the time-limited visual group, children were visually presented with each problem for the same length as stimuli in the auditory group (i.e., 3000 ms). Answers could be provided at any time. Note that the different presentation conditions included distinct combinations of presentation modality and presentation duration (see Table 1). In the analyses, we contrasted the groups with time-unlimited and time-limited visual presentation duration against each other and the groups with auditory and time-limited visual presentation modality against each other. Note that it was not possible to create a group with a time-unlimited auditory presentation that would be comparable to the time-unlimited visual presentation. That is, an auditory condition with artificially ongoing repeated stimulus presentation would have likely interfered with children's arithmetic processes. Sample sizes in third grade were 74, 67, 66 for the time-unlimited visual, time-limited visual and auditory groups, respectively; corresponding sample sizes in fourth grade were 49, 48, 43.

Table 1

Presentation conditions as combinations of presentation modality and duration.

		Presentation duration	
		Time-limited	Time-unlimited
Presentation modality	Visual	Time-limited visual group (reference condition; *n* = 115)	Time-unlimited visual group (*n* = 123)
	Auditory	Time-limited auditory group (*n* = 109)	–

Task procedure. Children worked on two blocks of problems with a brief resting break in-between blocks. Children were told that they should give an approximate answer for each addition problem without calculating the exact sum. They were instructed to use one of three rounding strategies, the rounding-down strategy (i.e., rounding both operands down to the nearest decades), the rounding-up strategy (i.e., rounding both operands up to the nearest decades), or the mixed-rounding strategy (i.e., rounding one operand up and the other down to the nearest decades). They were asked to give an estimate close to the exact sum while being as fast as possible.

Children were asked to indicate their answers on a touchscreen-based numpad. The numpad consisted of keys with rounded numbers from 10 to 200 to prevent children from giving exact sums or estimates with 5 as unit digit as answers. Five keys were aligned within 4 rows each, and numbers were arranged in ascending order from top left to bottom right. Prior to the computational estimation task, children entered eight verbally presented two-digit numbers (e.g., 60) on the numpad to ensure that they would be familiar with the input modality. Then, children practiced on six training problems with the presentation condition they were randomly assigned to. They were provided feedback on each practice problem to ensure that they would feel comfortable with executing the rounding strategies. No individuals had any difficulty with the rounding strategies and procedure.

Children's estimates were used to infer which strategy was used on each problem[7]. As an example, a child was considered to have used the rounding-down strategy when providing 90 as estimate for 41 + 58, the mixed-rounding strategy when providing 100, the rounding-up strategy when providing 110, and "strategy error" when providing any other response.

Working memory updating tasks. Working memory updating was assessed with four different tasks to obtain a stable and purer measure of updating. These four tasks were the spatial keep-track, day keep-track, frog-position updating, and colour updating tasks. In the *spatial keep-track* task (Dirk & Schmiedek, 2016), children were presented with either two or three differently coloured, fictitious creatures (i.e., monsters) at different positions on a 4 by 4 grid. Children were asked to remember the initial positions of the creatures. The creatures disappeared after 3000 ms. Then, arrows matching the creatures' colours were successively presented in the centre of the grid for 2500 ms with inter-stimulus-intervals of 500 ms. Each arrow indicated that the creature of the corresponding colour should move one block in the arrow's direction. Three arrows for two-creature trials and four arrows for three-creature trials were presented (i.e., three and four updating operations, respectively). Participants were asked to remember the new, updated positions of all creatures and to reproduce the final positions on an empty grid at the end of each trial. Each of the two difficulty levels (i.e., two vs. three creatures) was presented eight times.

[7] Children were additionally asked to indicate which strategy they used after each problem in one of the two test blocks. Analyses regarding children's best strategy selection for both strategy indicators obtained converging results (see Appendix for detailed comparison of results). Therefore, we can assume that results with inferred strategies from children's estimate were not distorted by systematic misclassifications but indeed resemble differences in strategy selection.

In a *day keep-track* task (Hammerstein et al., 2019), across 18 trials, participants were presented with either one, two, or three distinct calendar pages. On each page, the name for a weekday was written. Children were asked to remember the initial weekdays on the calendar pages. The names of the weekdays appeared for 3000 ms for one-calendar trials and for 4000 ms for two- and three-calendar trials. After displaying weekday names, a series of updating instructions (i.e., German words for *1 forward, 2 forward*, *1 backward* or *2 backward*, and corresponding arrows) was presented on empty calendar pages for 3000 ms, one at a time. Updating operations were presented with inter-stimulus-intervals of 500 ms. Participants were asked to remember the updated weekdays for each calendar and to reproduce the final weekdays. The weekdays had to be updated three times for one- and two-calendar and three or four times for three-calendar trials. Each of the three difficulty levels (i.e., one, two, or three-calendar trials) was presented six times.

Next, in the *frog-position updating* task (LeFevre et al., 2009), a frog moved across three to seven distinct positions on an array of eight irregularly positioned water lily leaves. Each position of the frog was displayed for 1750 ms with inter-stimulus-intervals of 250 ms. Children were asked to indicate either the last two, three, or four positions of the frog, in five trials each.

Finally, children worked on 16 trials of a *colour updating* task (Lee, Ng, Bull, Pe, & Ho, 2011). They were visually and verbally presented with sequences of four to eight colours, one at a time. Each colour was presented for 1500 ms before it was covered for 500 ms and followed by the presentation of the next colour. Children were asked to remember either the last two or three colours, in eight trials each. Stimuli consisted of eight colours with monosyllabic names.

The percentages of correctly remembered items in each of the four tasks was used as a measure of individuals' updating function. If more than half the trials of a difficulty level were missing for a participant due to task abortions or technical errors, the subtest score was not calculated. Thus, subtest scores were calculated for 345, 344, 344 and 345 children out of 347 children for the spatial keep-track, the day keep-track, the frog updating, and the colour updating tasks, respectively. The internal consistency of the updating tasks, as calculated by Cronbach's alpha for four balanced parcels each was acceptable to high (i.e., α = .93, .83, .85, .77, for the spatial keep-track, day keep-track, frog updating, colour updating tasks, respectively). To obtain a task-nonspecific updating value, the factor scores of a factor analysis with a single factor were used for further analyses. We centred individual's updating scores within grades to investigate the influence of updating beyond age-related effects; thus, a score of zero reflects the updating function of an average child within his or her grade.

Arithmetic fluency task. Children's arithmetic fluency was assessed independently of the target computational estimation task to control for baseline arithmetic fluency when examining effects of working memory updating. Children were presented with 24 two-digit addition problems with rounded decades (e.g., 20 + 50). They were asked to indicate as fast as possible the sum on a touchscreen-based numpad with the same design as in the computational estimation task. For the arithmetic fluency task, we ln-transformed reaction times on each trial to minimize skewness. Then, mean response times were calculated for each participant. Note that only response times of trials with correct responses were considered and children's response times that deviated more than 2.5 *SD* from their individual mean were excluded. Mean response times were *z*-standardized within grades and inverted; thus, a score of zero reflects the arithmetic fluency of an average child within his or her grade and higher values reflect better arithmetic fluency. The internal consistency, as calculated by Cronbach's alpha for four balanced parcels, was high with $\alpha = .90$.

Data Analysis. Children's strategy selection was examined with cross-classified multilevel models to allow for and to examine (a) between-participant variance and (b) between-item variance. The data have a two-level structure, with trials on Level 1 nested both within participants and within items on Level 2. Children's dichotomous strategy selection variable was analysed with logit models for binary responses. To be able to interpret effects in terms of probabilities that $y = 1$ (e.g., that children used the best strategy on a given problem), we calculated the predicted probabilities of children's strategy selection for the different values of predictors and report them in the following as percentages. We calculated the predicted probabilities as population-averaged probabilities by averaging over different simulated Level 2 random intercepts. Thus, percentages in the results can be interpreted as predicted probabilities that are true for the average of children and items with a certain predictor combination (see also Steele, 2009, for the calculation of predicted probabilities in logit models).

Results

Strategy variability. This analysis aimed at testing whether presentation conditions influenced children's tendency to decrease their variability in strategy use (i.e., leading them to restrict their strategy use to a single strategy more often). For this purpose, we calculated percentages of the most frequently used strategy for each participant and test block. The bimodal distribution of these scores (vertical histogram in Figure 1) points to the existence of two distinct subgroups of children: The first subgroup approaching the problems in a flexible manner and adjusting strategies on a problem-by-problem basis; the second subgroup adopted an inflexible approach, solving all or almost all problems with the same strategy. We created a dichotomous flexibility variable and ran a logistic regression on whether presentation conditions and individual differences influenced children's tendency to use strategies flexibly or restrict their strategy use to a single strategy. Children were more likely to approach test blocks in a flexible manner when problems were presented visually with no time limitation (PP= 77%; $\beta = 0.65$, 95% CI [0.05, 1.27], $p = .04$) and – surprisingly – when problems were presented auditorily (PP= 78%; $\beta = 0.71$, 95% CI [0.09, 1.35], $p = .03$) than when presented visually with time limitation (PP = 65%). Furthermore, children with more efficient updating were more likely to show flexible strategy selection than children with less efficient updating (average predicted probability for children with a standardized updating score of +1 *SD*, PP_{+1SD} = 82% vs. PP_{-1SD} = 64%; $\beta = 0.49$, 95% CI [0.23, 0.76], $p < .001$) and fourth graders were more likely to show flexible strategy selection than third graders (PP = 80% vs. PP = 68%; $\beta = 0.36$, 95% CI [0.10, 0.63], $p = .009$). Similarly, children with better arithmetic fluency were more likely to use a flexible approach than children with poorer arithmetic fluency (PP_{+1SD} = 80% vs. PP_{-1SD} = 65%; $\beta = 0.42$, 95% CI [0.16, 0.69], $p = .002$).

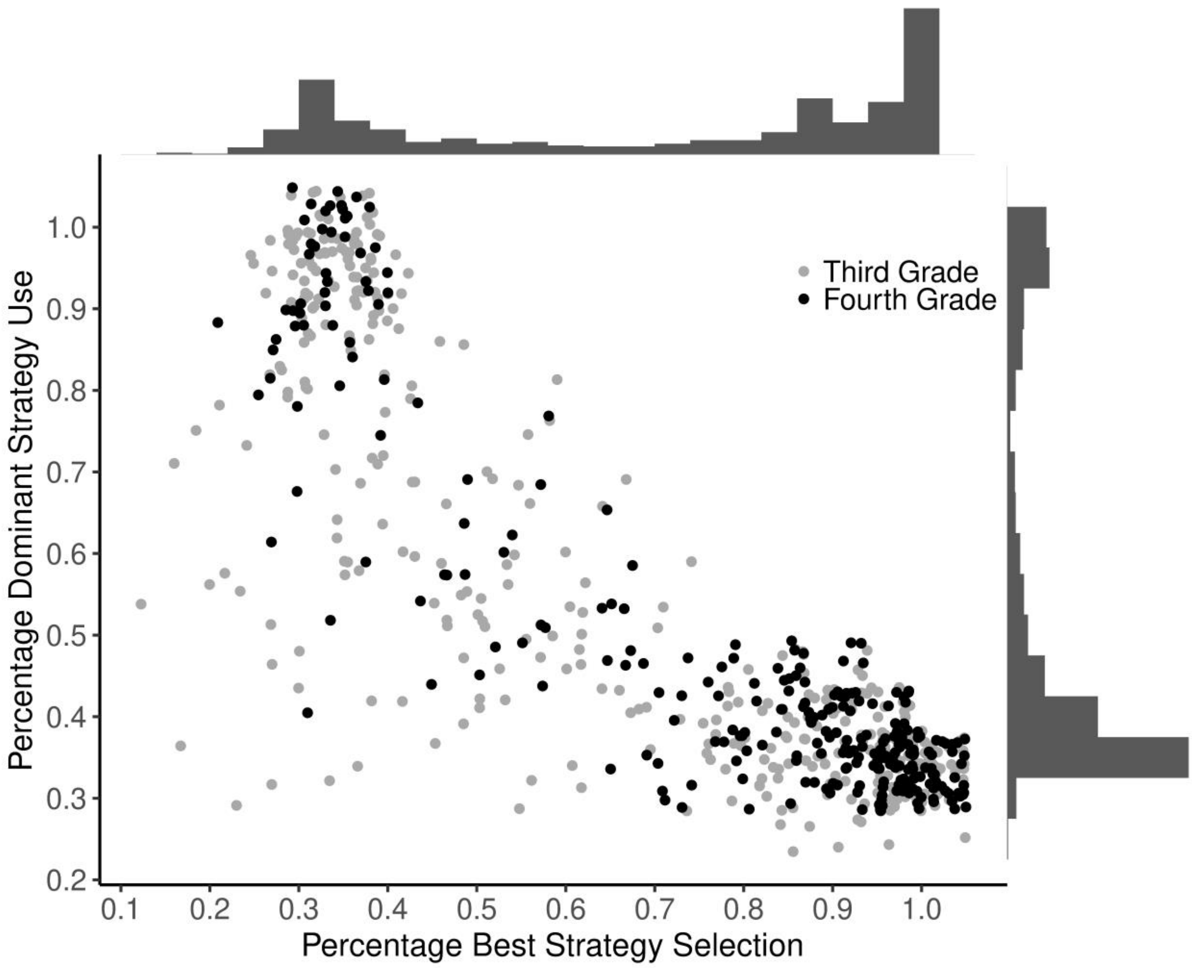


Figure 1. Relations between children's percentage dominant strategy use and percentage best strategy selection. Frequency distribution of percentages of dominant strategy use (vertical histogram) and percentages of best strategy selection (horizontal histogram).

Focus on test blocks with a flexible approach. Figure 1 revealed qualitatively different subgroups. Very high values in the percentage use of a dominant strategy (i.e., classification as inflexible) are (a) likely linked to an approach in which children do not select strategies on a trial-by-trial basis and (b) necessarily tied to best strategy selection of about 33%, whereas in flexible test blocks the adaptivity generally is clearly higher. Because analyses regarding children's strategy use would be distorted when including test blocks that were approached in an inflexible way, in the analysis of children's best strategy selection, we included only flexible test blocks with a dominant strategy being used on less than 75% of problems. This cut-off criterion was chosen as the distribution of the percentage use of the dominant strategy (vertical histogram in Figure 1) revealed a minimum at about 75%. Thus, the following analysis was run for a subsample of 277 children with data on at least one flexible test block (sample sizes in third grade were 47, 58, 51 for the time-unlimited visual, time-limited visual and auditory groups, respectively; corresponding sample sizes in fourth grade were 42, 41, 38).

Best strategy selection. The next analysis aimed at determining how presentation conditions influence children's best strategy selection and determine whether sensitivity to presentation conditions interacts with working memory updating. Answers were considered to be the best strategy and coded 1 if the inferred strategy matched the optimal problem-based strategy (i.e., the best strategy was considered the rounding-down strategy on small-unit problems, the mixed-rounding strategy on mixed-unit problems, and the rounding-up strategy on large-unit problems; see also Xu et al., 2014) and 0 otherwise. See Table 2 for a detailed list of predicted probabilities of children's best strategy selection.

As expected, children were more likely to choose the best strategy when problems were presented visually with no time limitation than when presented visually with time limitation (PP = 87% vs. PP = 79%; β = 1.20, 95% CI [0.70, 1.71], $p < .001$) or auditorily (PP = 73%). There was no significant difference in best strategy selection between time-limited visual and the auditory conditions ($p = .12$).

In addition, individual differences were found. Children with more efficient working memory updating were more likely to choose the best strategy than children with less efficient updating (PP_{+1SD} = 88% vs. PP_{-1SD} = 73%; β = 0.49, 95% CI [0.14, 0.84], $p = .006$) and fourth graders were more likely to choose the best strategy than third graders (PP = 84% vs. PP = 77%; β = 0.37, 95% CI [0.17, 0.57], $p < .001$). Furthermore, children with better arithmetic fluency were more likely to choose the best strategy than children with poorer arithmetic fluency (PP_{+1SD} = 83% vs. PP_{-1SD} = 77%; β = 0.25, 95% CI [0.03, 0.47], $p = .03$). Note that these main effects of individual differences occurred under the control of the other individual variables.

Interestingly, the interaction between presentation duration and updating was significant (β = 0.53, 95% CI [0.05, 1.01], $p = .03$). Contrary to our predictions, this occurred because children with more efficient updating benefitted even more from time-unlimited presentation relative to the time-limited condition (see Figure 2). The interaction between presentation modality and updating was not significant ($p = .42$).

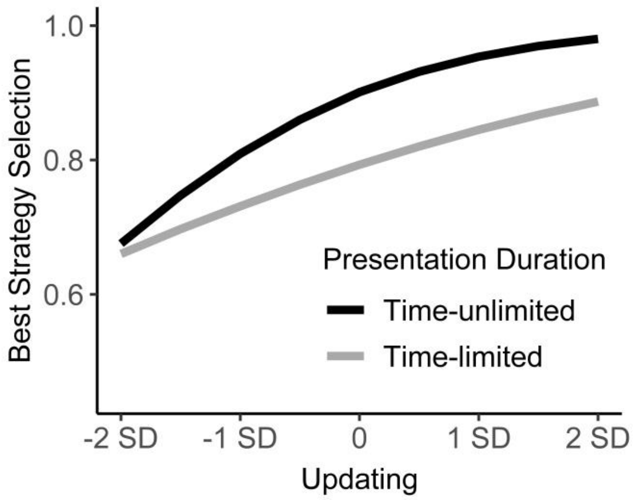


Figure 2. Influence of updating and presentation duration on predicted probabilities of children's best strategy selection.

Table 2

Predicted probabilities of better strategy selection (%) of third and fourth graders with varying levels of updating in the different presentation groups.

		Presentation group			
	Grade	Time-unlimited visual	Time-limited visual	Time-limited auditory	Mean
Less efficient updating (-1 *SD*)	Third	77	68	58	69
	Fourth	86	78	70	78
	Mean	81	73	64	73
More efficient updating (+1 *SD*)	Third	94	81	79	86
	Fourth	97	88	87	91
	Mean	95	85	83	88
Mean	Third	85	74	68	77
	Fourth	90	83	78	84
	Mean	87	79	73	80

Note. Unlike in an analysis of variance, children were not split into subsamples with less and more efficient updating, but the effect of updating was modelled continuously. For illustration, predicted probabilities are shown for children with updating processes 1 *SD* above average and 1 *SD* below average.

Discussion

Data revealed that strategic aspects of children's performance varied as a function of presentation conditions. Time-limited visual presentation led more children to use a single strategy on (almost) all problems. These results might be best explained by children adjusting their strategy use to the transience of visual stimuli: Under time-limited presentation, children might save time and cognitive resources by using a single strategy on (almost) all problems. By using the same strategy throughout the task and not adapting strategies to the problems, children do not have to encode the entire task but only the decade digits to yield a reasonable estimate, thereby alleviating the cognitive burden associated with strategy selection processes. In addition, this approach might enable children to accomplish some of the procedures involved in computational estimation within the 3 s of visual representation of the task. In contrast, children do not have a visual representation when problems are presented auditorily. This leads them to have to encode problems with precision, including processing the size of unit digits, a crucial feature for strategy selection processes. This might explain why children tested in the auditory condition showed greater strategy variability, even though presentation was time-limited in this group as well.

As another important finding, the presentation condition influenced flexible children's best strategy selection. As hypothesized, children were more likely to select the best strategy when problems were visually presented with no time limitation than when problems were visually presented with time limitation or auditorily presented. Benefits from time-unlimited visual presentation are best explained by the problem display serving as a permanent, external input (Adams & Hitch, 1997), which participants can refer to throughout the trials, enhancing the efficient encoding of problem features that are crucial for selecting the best strategy (e.g., size of units). Results on effects of presentation duration on strategy use in adults have been inconsistent with one study showing that shorter presentation moderately encouraged undergraduates to select the best strategy more often than longer presentation (Xu et al., 2014) and another study reporting better strategy selection for longer than shorter presentation (Lemaire & Brun, 2016). The present findings might indicate that for children as young as third and fourth graders strategy tasks impose task demands for which they need more time to solve them accurately. Therefore, an external representation of the task seems to be most beneficial for children this age.

Interestingly, children's best strategy selection did not differ between time-limited visual and auditory conditions. This indicates that the activation of different representational codes for auditory or visual input did not crucially interact with strategy selection mechanisms.

Instead, the presence of an external representation that children can refer to throughout the task increased their accuracy of strategy selection.

Importantly, findings showed interactions between updating and presentation conditions. In contrast to our prediction, children with more efficient updating benefitted from time-unlimited visual presentation to select the best strategy even more than children with less efficient updating. This indicates that these children used the advantages of a continuous problem display in a more efficient manner. This could occur via several mechanisms. For example, children might more flexibly change the focus of their attention by re-encoding the task when needed, or by checking the plausibility of their estimates more often than children with less efficient updating. Indeed, updating consists in "coding incoming information for relevance to the task at hand" (Miyake et al., 2000, p. 57). Therefore, children with more efficient updating might have had an advantage to identify problem characteristics crucial for selecting the best strategy.

Beyond individual differences, it is still important to emphasize that continuous presentation did support all children in finding estimates regardless of their updating capacities. Children with less efficient updating showed similar performance in selecting the best strategy when problems were presented visually with no time limitation (PP = 81%) as children with more efficient updating when presentation was auditory (PP = 83%) or visually with time limitation (PP = 85%). This indicates that supportive task environment (i.e., time-unlimited visual presentation) could help to circumvent cognitive disadvantages of children with less efficient updating to some extent.

Experiment 2: Strategy Execution

Experiment 1 showed that children used different strategies under varying presentation conditions. In this second experiment, we addressed the question whether children would execute strategies with different levels of efficacy in terms of accuracy and speed of execution.

Method

Participants. A total of 363 children were tested: The sample consisted of 227 third graders (121 males; age in months: *M* = 112.5, *SD* = 5.7; range = 99-133) and 136 fourth graders (76 males; age in months: *M* = 124.4, *SD* = 5.8; range = 113-143). Children were recruited from the same 50 classes as in Experiment 1, but no child participated in both experiments.

Procedure. We used the same procedure as in Experiment 1.

Material. Children worked on the same tasks as in Experiment 1 with differences in the computational estimation task.

Computational estimation task. Children worked on the same problems as in Experiment 1 and were randomly assigned to one of the three presentation groups. Sample sizes in third grade were 77, 73, 77 for the time-unlimited visual, time-limited visual and auditory groups, respectively; corresponding samples sizes in fourth grade were 42, 49, 45.

In contrast to Experiment 1, for each addition problem children were directed which strategy to use. The strategy was cued with two arrows presented besides the operands of the problem, with the direction of the arrows indicating how the operands had to be rounded (i.e., an arrow pointing down indicated that the operand had to be rounded down and an arrow pointing up indicated that the operand had to be rounded up). In the auditory condition, only the arrows but not the operands were visually presented. For each problem, the best strategy was cued. That is, the rounding-down, rounding-up, and mixed-rounding strategies were cued on the small-unit, large-unit, and mixed-unit problems, respectively.

Data Analysis. Children's strategy execution was examined with cross-classified multilevel models to allow for and to examine (a) between-participant variance and (b) between-item variance. The data have a two-level structure, with trials on Level 1 nested both within participants and within items on Level 2. We investigated whether different presentation conditions would be accompanied by different levels of efficacy of strategy execution in terms of (a) accuracy of execution (i.e., whether children's response matched estimates according to the cued strategy) and (b) speed of execution (i.e., estimation latencies). Children's accuracy of execution was examined with logit models for binary responses. Percentages in the results on accuracy of execution can be interpreted as predicted probabilities that are true for the average of children and items with a certain predictor combination (see also Steele, 2009, for the calculation of predicted probabilities in logit models).

Results

Accuracy of execution. This analysis aimed at determining whether children executed cued strategies accurately (i.e., whether responses matched estimates that would be yielded by executing the cued strategy). Answers were coded 1 if the response matched the estimate according to the cued strategy and 0 otherwise. See Table 3 for a detailed list of predicted probabilities of children's accurate strategy execution.

Results revealed that children were more likely to execute cued strategies accurately when problems were presented visually with no time limitation than when problems were presented visually with time limitation (PP = 88% vs. PP = 76%; β = -1.13, 95% CI [-1.53, -0.74], $p < .001$) or presented auditorily (PP = 69%). There were no significant differences in accuracy of execution between time-limited visual and auditory groups ($p = .11$).

Children with more efficient updating were more likely to execute cued strategies accurately than children with less efficient updating (PP_{+1SD} = 88% vs. PP_{-1SD} = 69%; β = -0.56, 95% CI [-0.86, -0.27], $p < .001$) and fourth graders were more likely to execute cued strategies accurately than third graders (PP = 84% vs. PP = 74%; β = -0.46, 95% CI [-0.63, -0.30], $p < .001$). Furthermore, children with better arithmetic fluency were more likely to execute cued strategies accurately than for children with poorer arithmetic fluency (PP_{+1SD} = 82% vs. PP_{-1SD} = 73%; β = -0.01, 95% CI [-0.02, -0.01], $p = .002$).

Interestingly, the effect of presentation modality was moderated by children's updating capacities (β = -0.53, 95% CI [-0.92, -0.14], $p = .009$) as effects of modality decreased as children's efficiency of updating increased (see Figure 3). That is, children with more efficient updating were equally accurate in executing strategies when problems were presented visually and auditorily whereas children with less efficient updating were less accurate in executing strategies when problems were auditorily presented. With very high levels of updating (>1 *SD* above average) children tended to be even more accurate when problems were presented auditorily than visually. This indicates that specifically children with less efficient updating benefitted from visual presentation regarding their accuracy of strategy execution.

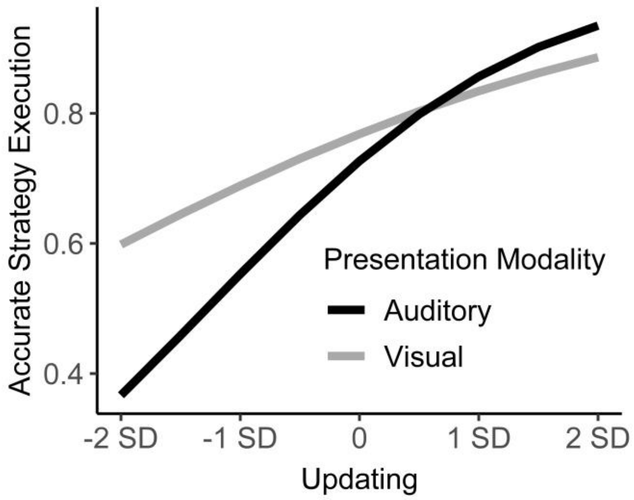


Figure 3. Influence of updating and presentation modality on predicted probabilities of children's accurate strategy execution.

Table 3

Predicted probabilities of accurate strategy execution (%) of third and fourth graders with varying levels of updating in the different presentation groups.

		Presentation group			
	Grade	Time-unlimited visual	Time-limited visual	Time-limited auditory	Mean
Less efficient updating (-1 *SD*)	Third	79	64	49	64
	Fourth	89	78	65	78
	Mean	83	69	55	69
More efficient updating (+1 *SD*)	Third	91	80	83	85
	Fourth	96	89	91	92
	Mean	93	83	86	88
Mean	Third	85	72	65	74
	Fourth	92	83	77	84
	Mean	88	76	69	78

Note. Unlike in an analysis of variance, children were not split into subsamples with less and more efficient updating, but the effect of updating was modelled continuously. For illustration, predicted probabilities are shown for children with updating processes 1 *SD* above average and 1 *SD* below average.

Speed of execution. This analysis aimed at determining whether children's speed of strategy execution (i.e., estimation latencies) were influenced by presentation and individual characteristics. See Table 4 for a detailed list of predicted estimation latencies of children's strategy execution.

Results showed that children were faster when problems were presented visually with time limitation (9 s) than when problems were presented visually with no time limitation (12 s; $\beta = 2.38$, 95% CI [1.34, 3.42], $p < .001$) or auditorily (14 s; $\beta = 5.06$, 95% CI [4.02, 6.10], $p < .001$).

Fourth graders were faster than third graders (10 vs. 13 s; $\beta = -1.24$, 95% CI [-1.68, -0.81], $p < .001$) and children with better arithmetic fluency were faster than children with poorer arithmetic fluency (10 vs. 14 s; $\beta = -2.02$, 95% CI [-2.47, -1.57], $p < .001$). The main effect of updating on children's estimation latencies was not significant ($p = .86$).

Interestingly, the effect of presentation modality was moderated by children's updating capacities ($\beta = -2.12$, 95% CI [-3.19, -1.06], $p < .001$) as effects of modality decreased as children's efficiency of updating increased. That is, children with more efficient updating were equally fast when problems were presented visually and auditorily whereas children with less efficient updating were slower when problems were auditorily presented (see Figure 4). This indicates that specifically children with less efficient updating benefitted from visual presentation regarding their estimation latencies.

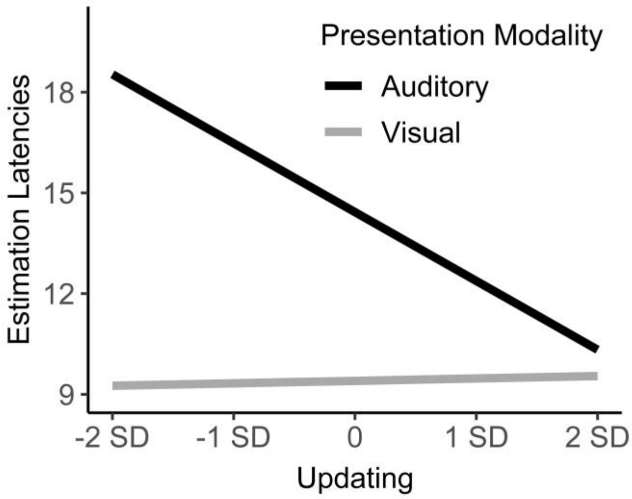


Figure 4. Influence of updating and presentation modality on children's predicted estimation latencies.

Table 4

Predicted estimation latencies (in seconds) of third and fourth graders with varying levels of updating in the different presentation groups.

		Presentation group			
	Grade	Time-unlimited visual	Time-limited visual	Time-limited auditory	Mean
Less efficient updating (-1 *SD*)	Third	13.1	10.2	17.4	13.6
	Fourth	10.6	7.8	15.0	11.2
	Mean	12.1	9.3	16.5	12.7
More efficient updating (+1 *SD*)	Third	12.3	10.4	13.3	12.0
	Fourth	9.8	7.9	10.9	9.6
	Mean	11.3	9.5	12.4	11.1
Mean	Third	12.7	10.3	15.4	10.4
	Fourth	10.2	7.8	12.9	12.8
	Mean	11.7	9.4	14.4	11.9

Note. Unlike in an analysis of variance, children were not split into subsamples with less and more efficient updating, but the effect of updating was modelled continuously. For illustration, predicted estimation latencies are shown for children with updating processes 1 *SD* above average and 1 *SD* below average.

Discussion

Effects of presentation conditions were not only seen on strategy selection (Expt. 1) but also found on strategy execution. That is, children were fastest when problems were visually presented with time limitation. Again, this might be explained in terms of the transience of visual stimuli in this condition leading children to try to be as fast as possible in estimating the sums. However, being fast came with costs: Children were less likely to execute strategies accurately under time-limited than under time-unlimited display. This is consistent with findings that children make more calculation errors when presentation is time-limited (Fürst & Hitch, 2000). These findings also imply that children are less accurate when they focus on speed (for similar results in adults, see Xu et al., 2014).

Time-limited visual and auditory conditions did not differ in terms of accuracy of strategy execution. Interestingly, however, children's working memory updating moderated effects of presentation modality on strategy execution. That is, children with more efficient updating were equally fast in the time-limited visual condition and the auditory condition whereas children with less efficient updating were slower under auditory condition. Similarly, children with more efficient updating were equally accurate in executing strategies in the time-limited visual condition and the auditory condition whereas children with less efficient updating were less accurate under auditory condition. This indicates that updating facilitated speed and accuracy of processing auditorily presented material. Hence, the present findings yield evidence that the activation of representational codes for auditory input interfered more with arithmetic processes, such as rounding and simple addition, than visual codes only in children with less efficient updating. Updating seems to help children to process and coordinate auditory material and consecutive mental arithmetic processes, likely as updating consists in keeping track of task components and actively manipulating contents in working memory (Lehto, 1996; Morris & Jones, 1990). Interestingly, prior studies reported more errors in adults solving simple multiplication tasks when presentation was auditorily than visually (Klingner et al., 2011; LeFevre et al., 2001). It is somewhat surprising to detect effects of presentation modality on mental arithmetic in adults but not in a subsample of children reported here (i.e., children with more efficient updating). This might be explained by the task difficulty being higher in the studies in adults. That is, presentation of visual and auditory stimuli was crucially shorter (about 1 s) in the studies by Klingner and colleagues (2011) and by LeFevre and colleagues (2001) than in the present study (3 s). Also, adults might not be as used to auditorily presented tasks and mental calculation whereas children are more often confronted with such tasks in educational contexts. Therefore, task difficulty of auditorily presented problems might have been relatively lower for children with more efficient updating working on the present tasks than for adults working on tasks in prior studies. However, it would be interesting to investigate whether effects of presentation modality in adults would also differ with individual differences such as working memory updating.

Finally, time-unlimited visual presentation was most beneficial for children to accurately execute strategies. Similar to results in Experiment 1, this is best explained by the permanent, external representation of the task (Adams & Hitch, 1997) facilitating children to deeply encode problem features which are crucial for accurate arithmetic processes and to refer back to the task if needed to control for calculation errors.

General Discussion

In two experiments, we examined the effects of presentation modality and presentation duration on children's strategy selection and performance while they accomplished a computational estimation task. Large samples of third and fourth graders with varying working memory updating capacities were asked to estimate sums for two-digit addition problems with one of the available rounding strategies. The present data showed that strategic aspects of children's performance varied as a function of presentation modality and duration and that these presentation effects were moderated by individual differences in working memory updating. The present findings have important theoretical and empirical implications to further understand strategic aspects of cognitive performance. We now discuss these implications.

We found that a permanent, external representation of the task (i.e., unlimited visual display) was most beneficial for strategy processes. Findings imply that an external representation of the problem continuously available throughout task execution helped children to encode the task more precisely not only to select the best strategy more often but also to execute strategies accurately during the calculation process. Updating facilitated best strategy selection processes specifically when an external representation was present. This might occur by children with more efficient updating re-encoding the task when needed, or by checking the plausibility of their estimates more often.

Limiting the presentation duration of visual tasks seems to impose time-pressure. Results indicate that the task-goal of children from the time-limited visual group shifted towards speed and simplicity of estimation rather than accuracy and proximity, leading them to adjust their strategic behaviour. That is, more children in the time-limited visual group opted for an inflexible approach to the task (i.e., using a single strategy on all or almost all problems), thereby disengaging from cognitive demands of selection processes, and were faster in executing strategies than children in the other groups.

Both time-limited visual and auditory presentation led children to choose the best strategy less often and executed strategies less accurately than when presentation was time-unlimited. Interestingly, updating facilitated speed and accuracy of processing auditorily presented material.

The present study was the first to investigate the influence of presentation duration and modality on children's strategy use in computational estimation. Results showed that a time-unlimited presentation of the task was most beneficial for strategy selection and accuracy of strategy execution. Thus, not only simple calculation processes were supported by an external aid (Adams & Hitch, 1997; Fürst & Hitch, 2000), but also processes involved in strategy selection. Interestingly, results indicate that updating facilitated efficient encoding processes of time-unlimited presented materials in strategy selection. Additionally, updating facilitated strategy execution of auditorily presented material. Note that we did not test children under a time-unlimited auditory condition as an artificially ongoing repeated stimulus presentation would have likely interfered with children's arithmetic processes. However, one could think of future studies enabling children to repeat auditory stimuli and comparing their strategy use to that in a task condition with visual stimuli that can be repeated.

The present findings have important implications for understanding processes underlying estimation (and more generally arithmetic) performance and strategy selection, as well as how some task parameters may help children to improve their academic performance in a variety of educational contexts. Data imply that children's strategy use is substantially influenced by the availability of an external representation of the task. Existing models of strategies (Lovett & Anderson, 1996; Rieskamp & Otto, 2006; Shrager & Siegler, 1998) should be broadened to make assumptions on the influence of different presentation parameters on participant's strategy selection. For example, models of strategy choices assume that strategy selection depends on individuals' expertise and relative strategy benefits. Many previous findings are consistent with this assumption. However, influence of presentation conditions and children's updating capacities on strategy choices found here (Expt. 1) suggest that strategy choices are also influenced by situation characteristics, like whether the problem is visually versus auditorily presented, with or with no time limitation. Additional assumptions within models of strategy selection may enable them to account for effects of such situational parameters on strategy selection. As another example, one important assumption within models of strategies is that relative efficiency of strategies depends on complexity of procedures. Influences found here on strategy execution (Expt. 2) suggest that relative complexity varies with situational factors, i.e. presentation conditions, and individuals' updating capacities.

Associated costs of accurately and efficiently executing strategies for auditory tasks might be relatively smaller with increasing efficiency of updating.

References

Adams, J. W., & Hitch, G. J. (1997). Working memory and children's mental addition. *Journal of Experimental Child Psychology*, *67*(1), 21–38. https://doi.org/10.1006/jecp.1997.2397

Bisanz, J., & LeFevre, J.-A. (2013). Strategic and nonstrategic processing in the development of mathematical cognition. In D. F. Bjorklund (Ed.), *Children's Strategies: Contemporary Views of Cognitive Development* (pp. 213–243). Psychology Press.

Bull, R., & Lee, K. (2014). Executive Functioning and Mathematics Achievement. *Child Development Perspectives*, *8*(1), 36–41. https://doi.org/10.1111/cdep.12059

Cohen Kadosh, R., & Dowker, A. (Eds.). (2015). *The Oxford handbook of numerical cognition*. Oxford University Press.

Dirk, J., & Schmiedek, F. (2016). Fluctuations in elementary school children's working memory performance in the school context. *Journal of Educational Psychology*, *108*(5), 722–739. https://doi.org/10.1037/edu0000076

Dowker, A. (1997). Young Children's Addition Estimates. *Mathematical Cognition*, *3*(2), 140–153. https://doi.org/10.1080/135467997387452

Dowker, A., Flood, A., Griffiths, H., Harriss, L., & Hook, L. (1997). Young Children's Addition Estimates. *Mathematical Cognition*, *3*(2), 140–153. https://doi.org/10.1080/135467997387452

Fürst, A. J., & Hitch, G. J. (2000). Separate roles for executive and phonological components of working memory in mental arithmetic. *Memory & Cognition*, *28*(5), 774–782. https://doi.org/10.3758/BF03198412

Gilmore, C., Göbel, S., & Inglis, M. (2018). *An introduction to mathematical cognition*. *International texts in developmental psychology*. Routledge.

Hammerstein, S., Poloczek, S., Lösche, P., Lemaire, P., & Büttner, G. (2019). Effects of working memory updating on children's arithmetic performance and strategy use: A study in computational estimation. *Journal of Experimental Child Psychology*, *184*, 174–191. https://doi.org/10.1016/j.jecp.2019.04.003

Klingner, J., Tversky, B., & Hanrahan, P. (2011). Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks. *Psychophysiology*, *48*(3), 323–332. https://doi.org/10.1111/j.1469-8986.2010.01069.x

Lee, K., Ng, S. F., Bull, R., Pe, M. L., & Ho, R. H. M. (2011). Are patterns important? An investigation of the relationships between proficiencies in patterns, computation, executive functioning, and algebraic word problems. *Journal of Educational Psychology*, *103*(2), 269–281. https://doi.org/10.1037/a0023068

LeFevre, J.-A., Greenham, S. L., & Waheed, N. (1993). The Development of Procedural and Conceptual Knowledge in Computational Estimation. *Cognition and Instruction*, *11*(2), 95–132. https://doi.org/10.1207/s1532690xci1102_1

LeFevre, J.-A., Lei, Q., Smith-Chant, B. L., & Mullins, D. B. (2001). Multiplication by eye and by ear for Chinese-speaking and English-speaking adults. *Canadian Journal of Experimental Psychology*, *55*(4), 277–284. https://doi.org/10.1037/h0087374

LeFevre, J.-A., Skwarchuk, S.-L., Smith-Chant, B. L., Fast, L., Kamawar, D., & Bisanz, J. (2009). Home numeracy experiences and children’s math performance in the early school years. *Canadian Journal of Behavioural Science*, *41*(2), 55–66. https://doi.org/10.1037/a0014532

Lehto, J. (1996). Are Executive Function Tests Dependent on Working Memory Capacity? *The Quarterly Journal of Experimental Psychology Section a*, *49*(1), 29–50. https://doi.org/10.1080/713755616

Lemaire, P. (Ed.). (2018). *Cognitive development from a strategy perspective: A Festschrift for Robert Siegler*. Routledge.

Lemaire, P., Arnaud, L., & Lecacheur, M. (2004). Adults' age-related differences in adaptivity of strategy choices: Evidence from computational estimation. *Psychology and Aging*, *19*(3), 467–481. https://doi.org/10.1037/0882-7974.19.3.467

Lemaire, P., & Brun, F. (2016). Effects of problem presentation durations on arithmetic strategies: a study in young and older adults. *Journal of Cognitive Psychology*, *28*(8), 909–922. https://doi.org/10.1080/20445911.2016.1215318

Lemaire, P., & Lecacheur, M. (2011). Age-related changes in children's executive functions and strategy selection: A study in computational estimation. *Cognitive Development*. https://doi.org/10.1016/j.cogdev.2011.01.002

Lovett, M. C., & Anderson, J. R. (1996). History of Success and Current Context in Problem Solving. *Cognitive Psychology*, *31*(2), 168–217. https://doi.org/10.1006/cogp.1996.0016

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

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology*, *81*(2), 111–121. https://doi.org/10.1111/j.2044-8295.1990.tb02349.x

Rieskamp, J., & Otto, P. E. (2006). Ssl: A theory of how people learn to select strategies. *Journal of Experimental Psychology. General*, *135*(2), 207–236. https://doi.org/10.1037/0096-3445.135.2.207

Shrager, J., & Siegler, R. S. (1998). SCADS: A Model of Children's Strategy Choices and Strategy Discoveries. *Psychological Science*, *9*(5), 405–410. https://doi.org/10.1111/1467-9280.00076

Siegler, R. S. (2007). Cognitive variability. *Developmental Science*, *10*(1), 104–109. https://doi.org/10.1111/j.1467-7687.2007.00571.x

Siegler, R. S., & Booth, J. L. (2005). Development of numerical estimation. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 197–212). Psychology Press.

Siegler, R. S., & Lemaire, P. (1997). Older and younger adults' strategy choices in multiplication: Testing predictions of ASCM using the choice/no-choice method. *Journal of Experimental Psychology. General*, *126*(1), 71–92. https://doi.org/10.1037/0096-3445.126.1.71

Sowder, J. T., & Wheeler, M. M. (1989). The Development of Concepts and Strategies Used in Computational Estimation. *Journal for Research in Mathematics Education*, *20*(2), 130. https://doi.org/10.2307/749278

Star, J. R., Rittle-Johnson, B., Lynch, K., & Perova, N. (2009). The role of prior knowledge in the development of strategy flexibility: the case of computational estimation. *ZDM*, *41*(5), 569–579. https://doi.org/10.1007/s11858-009-0181-9

Steele, F. (2009). *Module 7: Multilevel models for binary responses: Concepts*. Retrieved from: Centre for Multilevel Modelling, University of Bristol http://www.bristol.ac.uk/media-library/sites/cmm/migrated/documents/7-conceptssample.pdf.

Xu, C., Wells, E., LeFevre, J.-A., & Imbo, I. (2014). Strategic flexibility in computational estimation for Chinese- and Canadian-educated adults. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *40*(5), 1481–1497. https://doi.org/10.1037/a0037346

Supplementary Material

Comparison of Analyses of Children's Best Strategy Selection for the Two Distinct Strategy Indicators

We used two indicators to assess children's strategy selection. First, children's estimates were used to infer which strategy was used on each problem. Second, children were asked to indicate which strategy they used after each problem in one of the two test blocks. To do so, they pressed corresponding keys on the keyboard (i.e., one key each for each of the three available strategies and one key with a question mark in case children could not remember the just-executed strategy). Strategy query was randomly implemented in children's first or second test block.

We assessed children's strategy selection with two indicators to control for systematic misclassifications (e.g., misclassifications of inferred strategies due to calculation errors or misclassifications of strategy query due to unreliable self-reports). We ran analyses regarding children's best strategy selection for both strategy indicators as dependent variables and obtained converging results. Both measures yielded the same patterns of results (see Table S1). A minor difference was that the main effect of arithmetic fluency in the model of children's self-report was not significant as it was in the model of inferred strategies. Still, the marginal significance ($p = .052$) and narrow confidence interval support the robustness of the effect. Similarly, the interaction effect of updating and presentation duration was not significant in the second model. Again, the effect was marginally significant ($p = .08$). Therefore, we can conclude that results regarding children's best strategy selection are robust and were not distorted by systematic misclassifications but in fact resemble differences in children's strategy selection. We decided to report only results from inferred strategies in the manuscript's main body to be concise and to report results based on the complete set of trials and children assessed in the present study.

Table S1

Effects of the two models regarding children's best strategy selection for the different strategy indicators.

	Inferred Strategy (n_{trials} = 8359; $n_{participants}$ = 277)			Self-reports (n_{trials} = 4175; $n_{participants}$ = 255)		
Fixed Part	β	95% CI	*p*	β	95% CI	*p*
Intercept	1.96	[1.55, 2.37]	<.001	0.62	[0.33, 0.92]	<.001
Modality	-0.39	[-0.87, 0.10]	.12	0.14	[-0.24, 0.52]	.46
Duration	1.20	[0.70, 1.71]	<.001	0.93	[0.56, 1.29]	<.001
Updating	0.49	[0.14, 0.84]	.006	0.32	[0.06, 0.58]	.02
Grade	0.37	[0.17, 0.57]	<.001	0.22	[0.07, 0.37]	.005
Arithmetic Fluency	0.25	[0.03, 0.47]	.03	0.17	[-0.00, 0.34]	.052
Updating*Modality	0.22	[-0.32, 0.76]	.42	0.19	[-0.24, 0.61]	.40
Updating*Duration	0.53	[0.05, 1.01]	.03	0.29	[-0.04, 0.65]	.08
Random Part	*u*	Δ BIC		*u*	Δ BIC	
Participant intercept	2.18	1002.8		1.05	329.1	
Item intercept	0.44	226.8		0.18	61.6	

Note. Rows coloured dark grey contain results that match between the two models (i.e., effects of both models are not significant or are significant in the same direction); rows coloured light grey contain results that are significant in one model but not in the other while the slope is descriptively in the same direction; significant effects are indicated in boldface; Δ BIC = change in Bayesian Information Criterion, if random intercept dropped from model.

7 APPENDIX

A General Statements (in German)

Erklärung über frühere Promotionsversuche

Ich erkläre hiermit, dass ich mich bisher keiner Promotionsprüfung unterzogen habe.

Eidesstattliche Versicherung

Ich erkläre hiermit, dass ich die vorgelegte Dissertation mit dem Titel „Cognitive Processes in Children’s Strategy Use in Computational Estimation“ selbstständig angefertigt und mich anderer Hilfsmittel als der in ihr angegebenen nicht bedient habe, insbesondere alle Entlehnungen aus anderen Schriften mit Angabe der betreffenden Schrift gekennzeichnet sind. Ich versichere, die Grundsätze der guten wissenschaftlichen Praxis beachtet, und nicht die Hilfe einer kommerziellen Promotionsvermittlung in Anspruch genommen zu haben.

Erklärung zu Kriterien für publikationsbasierte Dissertationen

Ich erkläre nachfolgend, die Kriterien für publikationsbasierte Dissertationen zu erfüllen, die für den Fachbereich Psychologie und Sportwissenschaften der Goethe Universität Frankfurt gültig sind.

(1) Die publikationsbasierte Dissertation soll in der Regel 3 Schriften umfassen, die aus den letzten 5 Jahren stammen sollen.
Erklärung: Die Dissertation umfasst 4 Schriften aus den Jahren 2019 und 2020.

(2) Die Schriften sollen im Wesentlichen einem zusammenhängenden Forschungsprogramm entstammen. Die jeweils verfolgten Forschungsfragen sollen sich sinnvoll zueinander in Beziehung setzen lassen.
Erklärung: Die Schriften entstammen einem zusammenhängenden Forschungsprogramm zu kognitiven Prozessen bei der Anwendung von Rechenstrategien im Grundschulalter. Die vorliegende Arbeit enthält einen Text (siehe insbesondere die Kapitel *The Present Work* sowie *Connections Between the Studies*), in welchem die jeweils in den einzelnen Schriften verfolgten Forschungsfragen zueinander in Beziehung gesetzt werden.

(3) Der Kandidat oder die Kandidatin soll bei 2 Publikationen Erstautor/Erstautorin sein, bei einer weiteren Publikation kann er/sie Koautor/Koautorin sein. Eine geteilte Erstautorenschaft wird für jeden der Erstautoren anteilig gewichtet (bei 2 Erstautoren eine 1/2 Erstautorenschaft, bei 3 eine 1/3 Erstautorenschaft usw.).
Erklärung: Ich, Svenja Hammerstein, bin bei drei Publikationen Erstautorin sowie bei einer Publikation Koautorin.

(4) Die drei Schriften sollen zur Veröffentlichung zumindest eingereicht sein. Der aktuelle Status ist detailliert darzulegen (Publikationsorgan und Status wie eingereicht, in revision, conditional accept usw.).
Erklärung: Eine der vier Schriften ist veröffentlicht. Drei der vier Schriften sind zur Veröffentlichung eingereicht. Die erste Schrift wurde im August 2019 in der internationalen Fachzeitschrift „Journal of Experimental Child Psychology" veröffentlicht. Die zweite Schrift wurde am 31.07.2019 in der internationalen Fachzeitschrift „Journal of Numerical Cognition" eingereicht und am 23.09.2019 zur Revision eingeladen. Am 05.06.2020 wurde die revidierte Fassung erneut eingereicht

und ist unter Begutachtung. Die dritte Schrift wurde am 24.01.2020 in der internationalen Zeitschrift „Journal of Experimental Child Psychology“ eingereicht und am 28.03.2020 zur Revision eingeladen. Am 08.06.2020 wurde die revidierte Fassung erneut eingereicht und ist unter Begutachtung. Die vierte Schrift wurde am 08.06.2020 in der internationalen Fachzeitschrift „Quarterly Journal of Experimental Psychology“ eingereicht und ist unter Begutachtung.

(5) Mindestens 2 der 3 Schriften müssen in guten oder sehr guten, in der Regel englischsprachigen, Zeitschriften mit Peer-Review eingereicht sein.
Erklärung: Alle vier Schriften sind bei sehr guten, englischsprachigen Zeitschriften mit Peer-Review eingereicht oder veröffentlicht („Journal of Experimental Child Psychology“, „Journal of Numerical Cognition“ und „Quarterly Journal of Experimental Psychology“)

(6) Eine der 3 Schriften kann als Publikation in einem einschlägigen Lehrbuch, Enzyklopädieband oder einem anderen für das jeweilige Fach bedeutsame Publikationsorgan, jeweils mit Peer-Review, eingereicht oder veröffentlicht sein.
Erklärung: Keine der vier Schriften ist in einem Lehrbuch, Enzyklopädieband oder einem anderen für das jeweilige Fach bedeutsame Publikationsorgan, jeweils mit Peer-Review, eingereicht.

(7) Die als Dissertation vorgelegte Abhandlung soll über die zusammengestellten Publikationen hinaus einen zusätzlichen Text enthalten, in welchem eine kritische Einordnung der eigenen Publikationen aus einer übergeordneten Perspektive heraus vorgenommen wird. Dieser Text sollte einen Umfang von ca. 30 Seiten haben. Es sollen die Fragestellungen theoretisch entwickelt werden, die empirischen Arbeiten und ihre Ergebnisse so dargestellt werden, dass sie auch ohne Lesen der Einzelarbeiten nachvollziehbar sind und es soll eine Gesamtdiskussion enthalten, die die Fragestellungen beantwortet und den Erkenntnisgewinn der Arbeit herausstellt.
Erklärung: Die vorliegende Arbeit enthält einen entsprechenden Text.

B Author's Contributions (in German)

Erklärung über die Eigenleistung

Hammerstein, S., Poloczek, S., Lösche, P., Lemaire, P., & Büttner, G. (2019). Effects of working memory updating on children's arithmetic performance and strategy use: A study in computational estimation. *Journal of Experimental Child Psychology*, *184*, 174-191.

Svenja Hammerstein entwickelte in Rücksprache mit Gerhard Büttner, Sebastian Poloczek und Patrick Lösche die Fragestellungen und das Design der Studie. Svenja Hammerstein entwarf die Untersuchungsparadigmen und war gemeinsam mit Patrick Lösche für die Programmierung ebenjener verantwortlich. Svenja Hammerstein führte die Datenerhebung sowie die deskriptive und inferenzstatistische Datenanalyse mittels Mehrebenenanalysen durch. Ferner war sie als Erstautorin federführend für die Verfassung des Manuskripts verantwortlich. In engmaschiger Rücksprache überarbeiteten alle Mitautoren das Manuskript inhaltlich und sprachlich.

Poloczek, S., **Hammerstein, S.**, & Büttner, G. (2020). Children's mixed-rounding strategy use in computational estimation. *Manuscript submitted for publication*.

Sebastian Poloczek entwickelte in Rücksprache mit Gerhard Büttner die Fragestellungen und das Design der Studie. Svenja Hammerstein und Sebastian Poloczek waren für die Erstellung der Untersuchungsparadimen sowie die Programmierung ebenjener zuständig. Svenja Hammerstein bereitete die Rohdaten zur Datenanalyse vor und vollzog die deskriptive Datenanalyse. Sebastian Poloczek wertete die Daten inferenzstatistisch mittels Mehrebenenanalysen aus. Sebastian Poloczek war als Erstautor federführend für die Verfassung des Manuskripts verantwortlich. Svenja Hammerstein verfasste den Methodenteil. Svenja Hammerstein und Gerhard Büttner überarbeiteten das Manuskript inhaltlich und sprachlich und stimmten der eingereichten Version zu.

Hammerstein, S., Poloczek, S., Lösche, P., Lemaire, P., & Büttner, G. (2020). Two versus three available strategies in children's strategy selection in a computational estimation task. *Manuscript submitted for publication.*

Svenja Hammerstein entwickelte in Rücksprache mit Gerhard Büttner, Sebastian Poloczek und Patrick Lösche die Fragestellungen und das Design der Studie. Svenja Hammerstein entwarf die Untersuchungsparadigmen und war gemeinsam mit Patrick Lösche für die Programmierung ebenjener verantwortlich. Svenja Hammerstein führte die Datenerhebung sowie die deskriptive und inferenzstatistische Datenanalyse mittels Mehrebenenanalysen durch. Ferner war sie als Erstautorin federführend für die Verfassung des Manuskripts verantwortlich. In engmaschiger Rücksprache überarbeiteten alle Mitautoren das Manuskript inhaltlich und sprachlich und stimmten der eingereichten Version zu.

Hammerstein, S., Poloczek, S., Lösche, P., Lemaire, P., & Büttner, G. (2020). Effects of presentation duration and modality on children's strategy selection and performance: A study in computational estimation. *Manuscript submitted for publication.*

Svenja Hammerstein entwickelte in Rücksprache mit Gerhard Büttner, Sebastian Poloczek und Patrick Lösche die Fragestellungen und das Design der Studie. Svenja Hammerstein entwarf die Untersuchungsparadigmen und war gemeinsam mit Patrick Lösche für die Programmierung ebenjener verantwortlich. Svenja Hammerstein führte die Datenerhebung sowie die deskriptive und inferenzstatistische Datenanalyse mittels Mehrebenenanalysen durch. Ferner war sie als Erstautorin federführend für die Verfassung des Manuskripts verantwortlich. In engmaschiger Rücksprache überarbeiteten alle Mitautoren das Manuskript inhaltlich und sprachlich und stimmten der eingereichten Version zu.