The role of prior knowledge and need for cognition for the effectiveness of interleaved and blocked practice

Lea Nemeth[1] · Frank Lipowsky[1]

Abstract

Interleaved practice combined with comparison prompts can better foster students' adaptive use of subtraction strategies compared to blocked practice. It has not been previously investigated whether all students benefit equally from these teaching approaches. While interleaving subtraction tasks prompts students' attention to the different task characteristics triggering the use of specific subtraction strategies, blocked practice does not support students in detecting these differences. Thus, low-prior-knowledge students would benefit from interleaving rather than blocking as it guides them through the learning-relevant comparison processes. Because these comparison processes are cognitively demanding, students' need for cognition (NFC) could influence the effectiveness of interleaved practice. The present study investigates the role of students' prior knowledge and NFC for the effectiveness of interleaved and blocked practice. To this end, 236 German third-graders were randomly assigned to either an interleaved or blocked condition. Over 14 lessons, both groups were taught to use four number-based strategies and the written algorithm for solving subtraction problems. The interleaved learners were prompted to compare the strategies, while the blocked learners compared the adaptivity of one strategy for different mathematical tasks. A quadratic growth curve model showed that prior knowledge had a positive influence on students' development of adaptivity in the blocked but not in the interleaved condition. Students' NFC had a positive impact in the interleaved condition, while it had no influence in the blocked condition. However, the effects of prior knowledge and NFC did not differ significantly between the two conditions.

Keywords Interleaved practice · Primary school · Subtraction strategies · Adaptivity · Prior knowledge · Need for cognition

✉ Lea Nemeth
nemeth@uni-kassel.de

1 Department of Educational Science, University of Kassel, Nora-Platiel-Straße 1, 34109 Kassel, Germany

Introduction

In school mathematics, new concepts of one subject area are usually introduced and practiced in a blocked fashion, i.e., one after another (Rohrer et al., 2020). This also applies to the introduction of subtraction strategies in primary school and especially to the standard written algorithm. The standard written algorithm is usually introduced after practicing number-based strategies, which are rarely discussed in the classroom afterwards. This may explain why numerous studies show that students barely use number-based strategies after the standard written algorithm has been introduced. Instead, they almost exclusively use the standard written algorithm to solve subtraction problems, irrespective of the task characteristics (e.g., Hickendorff, 2020; Selter, 2001; Torbeyns et al., 2017). This lack in flexibility can lead to a non-adaptive use of different subtraction strategies, though the adaptive use of strategies, i.e., choosing efficient strategies based on task characteristics, is a meaningful competence for primary school students (e.g., Baroody & Dowker, 2003; Kilpatrick et al., 2001; National Council of Teachers of Mathematics (NCTM), 2000).

In contrast to the usual blocked learning approach in mathematics classrooms, in which exemplars or tasks of different categories are presented one after another (e.g., AAABBBCCCC), interleaved practice intermixes tasks (e.g., ABCBCACBA; Richter et al., 2022b). Research has shown that this approach can promote students' adaptive use of subtraction strategies (Nemeth et al., 2021). The advantage of interleaved over blocked practice is often traced back to implicit comparison processes (Birnbaum et al., 2013), which explains why interleaved practice might be suitable to foster students' adaptive use of subtraction strategies: by using different subtraction strategies alternately, students are made aware of the differences between the strategies, that is, for which kinds of subtraction problems which strategy can be used adaptively.

However, students do not automatically compare if they are not explicitly prompted to do so (Durkin et al., 2017). Previous research shows that interleaved practice combined with explicit prompts to compare different tasks and solution methods can better foster students' accuracy when solving mathematical problems, reduce misconceptions, and lead to more flexible and adaptive use of strategies compared to blocked learning (e.g., Nemeth et al., 2019, 2021; Ziegler & Stern, 2014, 2016), though it is unclear whether all students benefit equally from this teaching approach. The comparison processes evoked by interleaved practice are cognitively demanding. At the same time, they support the learning of relevant processes. Thus, the current study investigates the role of students' prior knowledge and need for cognition (NFC), i.e., their engagement in and enjoyment of cognitive activities, for the effectiveness of interleaved practice combined with prompts to compare.

Adaptive use of subtraction strategies

There is a broad consensus among mathematics researchers and educators that students should be able to use strategies adaptively to solve problems (e.g., Baroody & Dowker, 2003; Hickendorff et al., 2022; Kilpatrick et al., 2001; NCTM, 2000). Using strategies adaptively represents a part of individuals' cognitive variability, which leads to faster and more accurate problem solving (Heinze et al., 2009b; Verschaffel et al., 1998). The definition of strategy proficiency by Lemaire and Siegler (1995) illustrates that strategy competence is more than calculating correctly: strategy proficiency comprises strategy repertoire (the repertoire of strategies one uses), strategy distribution (the frequency with which

Table 1 Overview of selected number-based strategies and the standard written algorithm

Decomposition strategies	
Stepwise strategy	Split strategy
6 5 4 − 3 2 8 = 3 2 6	7 5 6 − 4 2 3 = 3 3 3
6 5 4 − 3 0 0 = 3 5 4	7 0 0 − 4 0 0 = 3 0 0
3 5 4 − 2 0 = 3 3 4	5 0 − 2 0 = 3 0
3 3 4 − 8 = 3 2 6	6 − 3 = 3

Shortcut strategies	
Compensation strategy	Indirect addition
5 4 7 − 3 9 9 = 1 4 8	4 5 2 − 4 4 9 = 3
5 4 7 − 4 0 0 = 1 4 7	4 4 9 + 3 = 4 5 2
1 4 7 + 1 = 1 4 8	

Standard written algorithm (turning the counter-method; Selter et al., 2012)

$$
\begin{array}{r}
7\ 2\ 5 \\
-\ 4\ 5\ 3 \\
\hline
{}^{1} \\
\hline
2\ 7\ 2
\end{array}
$$

the strategies are used), strategy efficiency (speed and accuracy), and strategy adaptivity (choosing an appropriate strategy based on task characteristics or based on the individual efficiency when applying specific strategies). Star (2005) describes adaptivity[1] as deep procedural knowledge, which "is associated with comprehension, flexibility, and critical judgement" (p. 408). Thus, the adaptive use of strategies exceeds procedural knowledge, i.e., the ability to execute procedures (Hiebert & Lefevre, 1986). In our study, we take a normative perspective on adaptivity following several other studies (e.g., Blöte et al., 2000; Heinze et al., 2018; Torbeyns et al., 2009) and define adaptivity as the fit between a specific subtraction task and the subtraction strategy used (without considering the accuracy of students' solutions).

Number-based strategies are based on the specific characteristics of the task and one's knowledge of the number system and operations (Torbeyns & Verschaffel, 2016). While these strategies are typically performed mentally, students often note solution

---

[1] Star (2005) uses the term flexibility, which is equivalent to our understanding of adaptivity, i.e., to use a solution strategy that best fits the target task.

steps or interim results to relieve their working memory. When using *decomposition strategies* (stepwise strategy and split strategy, Table 1), the hundreds, tens, and ones of one or both numbers are decomposed before starting to solve the task. When using *shortcut strategies* (compensation strategy and indirect addition, Table 1), one must adapt the numbers and operations flexibly to task characteristics (Torbeyns et al., 2009). Despite the range of available strategies, numerous studies have shown that primary school students rarely solve subtraction problems adaptively, relying instead on a few standard approaches (e.g., Blöte et al., 2000; Heinze et al., 2009a; Hickendorff, 2020; Selter, 2001). For example, shortcut strategies are rarely used if they have not been taught systematically (De Smedt et al., 2010; Torbeyns et al., 2009; Van der Auwera et al., 2023), although they can accelerate problem-solving and reduce mental effort if used for appropriate tasks. This non-adaptive use of strategies becomes even more pronounced after the introduction of the standard written algorithm (Table 1)—a fixed step-by-step procedure for solving mathematical tasks (Torbeyns & Verschaffel, 2016)—as students often use it by default (e.g., Selter, 2001; Torbeyns & Verschaffel, 2016; Torbeyns et al., 2017).

Empirical findings show that explicit teacher-led instruction on how and when to use different subtraction strategies adaptively promotes students' use of shortcut strategies and enhances their adaptivity (e.g., De Smedt et al., 2010; Heinze et al., 2018; Nemeth et al., 2019, 2021; for an overview, see Heinze et al., 2020). To foster students' adaptive use of strategies, it seems worthwhile to teach different subtraction strategies in an interleaved fashion, thereby encouraging students to reflect on their strategy choice for different tasks.

Interleaved practice and comparison learning

Despite the fact that teachers usually try to simplify learning for their students, there is large empirical evidence that hampering learning processes can lead to better long-term retention (Dunlosky et al., 2013). These so-called *desirable difficulties* include, among other instructional approaches, interleaved practice (Bjork & Bjork, 2011; Richter et al., 2022a). While blocked practice facilitates *within-comparisons*, which can help students to recognize common features of specific categories, it does not encourage students to compare *between* categories; they usually learn one category after another without comparing (Rohrer et al., 2015). Regarding solution strategies in mathematics, this may result in a lack of understanding of the underlying principles and the specific application conditions of each strategy (Ziegler et al., 2018).

Unlike blocked practice, interleaved practice intermixes learning contents, which hampers learning in the short-term. However, interleaving contents results in better long-term retention, as various studies demonstrate (Dunlosky et al., 2013). One theoretical explanation for this advantage is the discriminative-contrast hypothesis, which states that interleaved practice activates discrimination and comparison processes (Birnbaum et al., 2013). As learners are alternately confronted with different content, they are stimulated to focus on the differences (*attentional bias framework*; Carvalho & Goldstone, 2015, 2017). Moreover, students must choose a strategy for each problem and are required to reflect on their strategy choice which should benefit their adaptivity. When solving tasks in a blocked fashion, on the contrary, students are likely to assume that the current problem can be solved adaptively with the same strategy used for the previous task (Rohrer & Hartwig, 2020). Therefore, blocked practice does not engage students in reflecting their strategy choice to the same extent as interleaved practice.

Brunmair and Richter (2019) found a moderate interleaving effect in their meta-analysis over all included primary studies ($g = 0.42$) and a small interleaving effect for mathematics ($g = 0.34$). However, the included primary studies on interleaving mathematical tasks are inconsistent, yielding both negative and positive effects. The results of this meta-analysis suggest that the concrete implementation of interleaved practice influences its effectiveness. According to the discriminative-contrast hypothesis, comparison processes are expected to be underlying learning mechanisms, explaining the advantage of interleaved over blocked practice. Comparing supports students in learning the principles of each category and in detecting central differences and similarities among them (e.g., Gentner, 1983; Loewenstein et al., 1999). In the case of younger children, empirical studies have shown that they are less likely to focus on the relevant dimension during learning (Cook & Odom, 1992; Thompson & Markson, 1998). However, interleaved practice requires learners to pay attention to the relevant alternating dimension to use the evoked comparison processes. What is more, previous research demonstrates that learning which supports students' comparison processes by including explicit comparison prompts results in higher learning gains in contrast to only offering opportunities to compare (e.g., Alfieri et al., 2013; Catrambone & Holyoak, 1989; Gentner et al., 2003). Further studies have demonstrated that prompting students to compare different types of tasks or solution strategies can enhance flexibility and adaptivity when learning mathematics (e.g., Durkin et al., 2023; Rittle-Johnson & Star, 2007, 2009). This is supported by research showing that interleaved practice combined with prompts to compare can foster students' adaptive strategy use (Nemeth et al., 2019, 2021), though it is unclear whether all students benefit equally from this teaching approach.

Prior knowledge and NFC as potential moderators of interleaved practice

Prior arithmetical knowledge affects the adaptive use of subtraction strategies among primary school students (Nemeth et al., 2019; Torbeyns & Verschaffel, 2016). Choosing an adaptive subtraction strategy for a subtraction problem requires demanding cognitive processes: students have to analyze the numbers in the subtraction task and choose an appropriate strategy from their strategy repertoire or invent a strategy appropriate to the task's characteristics. To meet these challenges, students need a conceptual understanding of numbers, as well as a grasp of the number system and arithmetic operations. As interleaved practice is a *desirable difficulty*, it is expected that prior knowledge might be even more relevant when interleaved practice and comparison processes are included in the learning process. Blocking learning contents, and thus practicing the same procedure repeatedly, could reduce the demands on working memory. In contrast, when interleaving and comparing subtraction strategies, more interacting elements need to be processed simultaneously, which increases cognitive load (Sweller & Chandler, 1994). Interleaving and comparing various subtraction strategies might thus become an undesirable difficulty when solving subtraction tasks with a given strategy is already challenging for a student (McDaniel & Butler, 2011). Nonetheless, interleaving subtraction strategies can trigger learning mechanisms central to the acquisition of adaptive, task-based use: students are repeatedly prompted indirectly—and in our study also directly through comparison prompts—to compare strategies for different tasks. Compared to blocked learning, this approach should support students more in abstracting the conditions of the application of different subtraction strategies which could especially benefit low-prior-knowledge students.

Following these two contradictory explanations, empirical findings regarding the role of prior knowledge for interleaved practice (e.g., Rau et al., 2010, 2014) and comparison learning in mathematics (Guo et al., 2012, 2014; Rittle-Johnson et al., 2009, 2012; Star & Rittle-Johnson, 2009) are also inconsistent. Researchers, therefore, conclude that the target knowledge type and the aspects of the learning content that are critical for student learning might play a major role if interleaving and comparing are beneficial for low-prior-knowledge students (Guo et al., 2012, 2014; Rau et al., 2014). According to variation theory (Kullberg et al., 2017; Marton & Booth, 1997; Marton & Pang, 2006), learning occurs when individuals discern and focus on the critical aspects of the target phenomenon. To achieve this, the learning content must vary in the specific dimension critical to the target learning goal (Marton & Pang, 2006). When the adaptive use of subtraction strategies is the target learning goal, identifying the characteristics of subtraction tasks, and knowing when to apply which strategy efficiently is critical for learning. To support students in discerning and focusing on the critical aspects, they should be encouraged to contrast different examples, i.e., to compare the efficiency of different subtraction strategies for different subtraction tasks.

Contrasting and comparing are inherent to the learning process regarding the adaptive use of subtraction strategies: students need to analyze the task and weigh and compare the adaptivity of the different strategies to choose the most adaptive one. Interleaved practice and comparing can stimulate these processes repeatedly and may support students in acquiring this kind of deep procedural knowledge (Star, 2005). In particular, it can address aspects critical to the learning of students with lower prior knowledge, i.e., the characteristics of the tasks are systematically varied, and students are repeatedly confronted with the adaptivity of different strategies. High-prior-knowledge students, on the other hand, might already be equipped to compare and abstract the rules over longer periods, and therefore may also benefit from blocked practice (Rau et al., 2014).

NFC refers to an individual's intrinsic cognitive motivation, i.e., their engagement in and enjoyment of cognitive activities (Cacioppo et al., 1996). Individuals high in NFC process information more deeply, choose strategies more adaptively, and have a more positive attitude towards cognitively challenging tasks (Cacioppo et al., 1996; Evans et al., 2003), which explains why NFC has been associated with academic achievement (von Stumm & Ackermann, 2013), though in studies conducted with primary students, no relation was found (Ginet et al., 2000; Luong et al., 2017). However, there is still a lack of research regarding the impact of NFC when it comes to cognitively demanding tasks in the context of interleaved practice or comparison learning. As mentioned above, interleaved practice, as well as comparing learning contents, require higher cognitive effort from students compared to a blocked or sequential approach. Thus, it can be assumed that low-NFC students may not exploit the learning opportunities offered by these cognitively demanding teaching approaches. Hence, students' NFC may influence the effectiveness of interleaved practice combined with prompts to compare, while it may be less important for blocked practice, which is less cognitively demanding.

The current study

Several studies show that interleaved practice leads to higher learning gains than blocked practice (Brunmair & Richter, 2019). This is also true in primary school mathematics

Table 2 Overview of student characteristics for the interleaved and the blocked condition

Variable	Interleaved practice	Blocked practice
Age	9.04 (0.40)	9.10 (0.43)
Female	45.38%	45.30%
Prior arithmetical knowledge	− 0.04 (1.42)	0.03 (1.51)

(Nemeth et al., 2019, 2021; Taylor & Rohrer, 2010). However, research investigating whether all students benefit equally from interleaved practice in the primary mathematics classroom is lacking. Thus, the goal of this study is to investigate whether the effectiveness of interleaved practice combined with prompts to compare is moderated by students' prior knowledge and NFC.

In the first step, we investigate if students' prior knowledge moderates the effect of interleaved vs blocked practice (*research question 1*). Comparing different subtraction strategies for several tasks is critical for learning how to choose adaptive strategies based on task characteristics. Especially low-prior-knowledge students should benefit from interleaved practice including prompts to compare. This is because the critical aspects are explicitly varied, and students are encouraged to contrast and compare, and thus to abstract rules as to when to apply which subtraction strategy. Blocked practice does not explicitly offer these comparison processes and does not directly support students' focus on and discernment of different task characteristics triggering the use of different subtraction strategies. Students probably need more learning-relevant prior knowledge to abstract the rules without being repeatedly instructed to compare—implicitly through interleaved practice and explicitly through comparison prompts. Therefore, we hypothesize that prior knowledge has a stronger positive effect on students' learning gains in blocked than in interleaved practice (H1).

We further investigate if students' NFC has a moderating effect (*research question 2*). It can be assumed that the higher the students' NFC, the more they benefit from interleaved practice as a desirable difficulty in learning. This is because they enjoy cognitively demanding activities, leading to deeper information processing. Even though there are currently no studies which investigate the role of NFC for the effectiveness of interleaved practice/comparison learning, we expect that high-NFC students prefer to use the learning opportunities offered by interleaved practice, including comparison prompts. At the same time, NFC should be less important for blocked practice as it is less cognitively demanding. Hence, we expect that students' NFC has a stronger positive impact on students' learning gains in interleaved compared to blocked practice (H2).

Material and methods

Participants

The research questions were answered with an experimental classroom study in which 236 German third-graders from 12 classes participated. The students were randomly assigned to either the blocked condition including within-comparison prompts ($n=117$) or the

interleaved condition including between-comparison prompts ($n = 119$). The regular class composition was broken up, and new learning groups were formed. The halves of two classes of one school that were randomly assigned to the interleaved condition were combined into the interleaved learning group, and the halves of two classes that were assigned to the blocked condition formed the blocked learning group. The students were not told that the lessons differed between the groups. To participate in this study, addition up to 1000 had to have been introduced in class before the intervention, while subtraction up to 1000 need not have been introduced. Table 2 shows the descriptive statistics of student characteristics for the blocked and interleaved condition. Analyses revealed no significant difference between the two groups with respect to students' age, $t(231) = 0.80$, $p = 0.43$, the proportion of female and male students, $\chi(1) = 0.00$, $p = 0.99$, and prior arithmetical knowledge, $t(219) = -0.36$, $p = 0.72$.

Design

The experimental study follows a 2×4 design: (group: interleaved vs blocked) $\times$ (time: T1: before the intervention, T2: 1 day later, T3: 1 week later, and T4: 5 weeks later; Fig. 1).

At T0, prior arithmetical knowledge and NFC were measured. The dependent variable, adaptivity, was measured immediately before the intervention (T1), immediately after the intervention (T2), and at two follow-up tests, i.e., 1 week (T3) and 5 weeks (T4) after the intervention.

The lessons were conducted by four trained staff members using precise lesson scripts for each of the 14 lessons, including teacher explanations and questions, scenarios of potential student behavior during whole-class discussions and individual work, respective teacher reactions (e.g., how to deal with incorrect student answers), and the time expected to be spent on each activity (for examples of a lesson script from the blocked condition, see Supplementary file 1; for the interleaved condition, see Supplementary file 2). The staff members had studied mathematical didactics for primary schools at university for at least six semesters. To be able to disentangle teacher effects from treatment effects, all staff members taught both conditions equally often. During the intervention, no regular mathematics lesson was held, and the students were not assigned any mathematics homework. They were not allowed to take the material home.
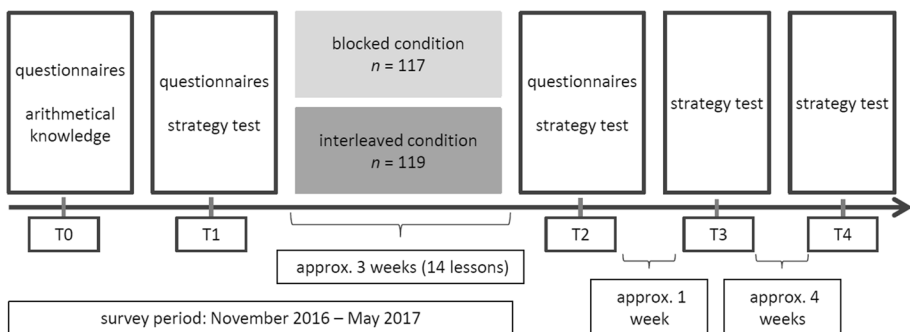

Fig. 1 Design of the study

Treatment

The intervention comprised 14 45-min lessons in which the adaptive use of subtraction strategies when solving three-digit subtraction problems was taught. The first two lessons were equal for both conditions: students' prior knowledge of numbers was activated and a first approximation of their aptitude at solving subtraction tasks was made at a math conference, i.e., groups of students discussed which strategy is the most appropriate for solving a specific subtraction task. Moreover, the criteria for solving tasks adaptively (number of solution steps, mental effort, error rate) were discussed in both conditions and a poster with these criteria was hung up in class in all following lessons (Supplementary file 3). From the third lesson onwards, the number-based subtraction strategies split strategy, stepwise strategy, compensation strategy, and indirect addition, along with the standard written algorithm (Table 1), were introduced and practiced in class. While the teaching content was the same, the two conditions differed in the order in which the strategies were introduced and practiced. In the blocked condition, the strategies were introduced and practiced one after another, while they were practiced alternately in the interleaved condition (detailed overview of the activities of each lesson for both conditions, see Supplementary file 4). The total time spent on each strategy and the mathematical tasks were nearly identical.

Posters with worked examples of each strategy were hung on the walls during the lessons to assist the students with calculations and with arguing whether a specific strategy is adaptive for a specific task (for an example, see Supplementary file 3). During the lessons, the teacher encouraged the students to use the posters to describe the solution steps of the different strategies (see Supplementary file 1 and 2). In addition, a poster (mathematical lexical storage) with relevant mathematical terms and the corresponding explanations was hung up in class and each student received a printout in A4 format to help the students verbalize explanations and solution steps (Supplementary file 3).

The students in the interleaved condition were explicitly prompted to compare strategies for specific tasks and to explain why a specific strategy is more adaptive than another (between-comparison) during individual work (see Fig. 2) and classroom discussions (e.g., "Which strategy is the cleverest for the task 441 – 297? The squirrel-strategy (compensation strategy), the mouse-strategy (stepwise strategy), or the frog-strategy (indirect addition)? Why do you think that this task can be solved cleverly with the squirrel-/mouse-/frog-strategy?"). In the blocked condition, each lesson focused on one strategy and the students in the blocked condition were not prompted to draw comparisons between strategies. Nevertheless, the specific task characteristics that prompt the use of each subtraction strategy were discussed and the students were asked to decide for several tasks if they can be solved adaptively with that strategy and to explain their decision (within-comparison, e.g., "Can you solve the task 441 – 297 cleverly with the squirrel-strategy or not? Why do you think that you can/cannot solve this task cleverly with the squirrel-strategy?"). Figure 2 illustrates the differences between the two conditions in individual work:

Fig. 2 Example worksheet of the blocked condition (left, lesson 7) and the interleaved condition (right, lesson 8). Squirrel-strategy = compensation strategy, mouse-strategy = stepwise strategy, frog-strategy = indirect addition. The tasks were translated from German to English

In this example, the students in the blocked condition had to decide for task 532 – 297 (and for several more tasks afterwards) if it can be solved adaptively with the squirrel-strategy (compensation strategy) and, if so, to solve the task with this strategy. The students in the interleaved condition, on the contrary, had to decide for the same task if the squirrel-, mouse-, or frog-strategy, i.e., compensation strategy, stepwise strategy, or indirect addition, is the most adaptive strategy and to solve the task with the most adaptive one.

To ensure that the students do not develop misconceptions, incorrect answers were corrected during whole-class discussion in both conditions. The students checked their answers during individual work with a solution sheet, which was only handed out when the students had solved all the tasks, or the answers were discussed in class (see Supplementary file 1 and 2 for examples).

Measures

Adaptivity

A subtraction strategy test was administered at T1–T4 to assess students' adaptivity. At each point, the test contained 11 items, of which six were linked across all points of measurement, while the other five items varied (anchor items: 532 – 476, 720 – 269, 534 – 399, 502 – 299, 802 – 797, 475 – 469). Students solved three-digit subtraction problems (exception: two two-digit tasks in the pretest). The tasks evoked the use of number-based strategies and the standard written algorithm. The students were prompted, "Solve the tasks in a clever way. Write down how you solved the tasks." The test time was 28 min to ensure that the students had sufficient time to reflect on the most adaptive strategy.

For each of the students' solutions, the used strategies were coded using a detailed coding system comprising 32 subtraction strategies[2] ($\kappa \geq 0.88$). Moreover, the adaptivity of students' solutions was assessed. Two independent raters estimated the adaptivity for each of the 32 strategies on a 3-point scale for each subtraction task with a standardized coding manual (0 = non-adaptive, 1 = partially adaptive, 2 = highly adaptive; Heinze et al., 2018). The following criteria were considered: number of solution steps, mental effort, and error rate (for an example rating, see Supplementary file 5). This rating was independent of the accuracy of the solutions: even if a student solved a subtraction task incorrectly due to a calculation error or inaccurate use of the strategy, the chosen strategy could still be rated as partially or highly adaptive (Heinze et al., 2018). Interrater reliability was calculated for each of the 26 different tasks over the ratings for each of the 32 strategies of both raters and was satisfactory overall ($0.63 \leq \kappa_{weighted} \leq 1.00$, $M = 0.87$, $SD = 0.10$). When raters did not agree, a consensus was negotiated. Adaptivity was scaled longitudinally over the four points of measurement using a one-dimensional partial credit model with virtual persons with ConQuest parametrization (Robitzsch et al., 2018). The advantage of using virtual persons to estimate person-parameters longitudinally is that the item difficulties are equated over the different points of measurement. The person-parameters of all points of measurement can be interpreted as values from the same scale. Therefore, the difference of person-parameters between different time-points can be interpreted as the growth or decline of individuals' abilities (Hartig & Kühnbach, 2006).

Four of the anchor items and nine of the varying items were dichotomized because the intermediate category was represented in less than 10%, whereas the other items remained trichotomous. One anchor item and one of the varying items of T3 were underfitting and thus removed from the model. The remaining items had an acceptable item fit ($0.67 < \text{WMNSQ} < 1.32$). The EAP/PV-reliability of 0.82, as well as the WLE-reliability of 0.79, were satisfactory ($\sigma^2_{WLE} = 1.92$). Recursive partitioning for partial credit models did not detect DIF between the two groups (Zeileis et al., 2018). The person-parameters were estimated using WLE. Person-parameters do not have an explicit minimum or maximum, and they can reach negative as well as positive values, with higher values representing a higher level of adaptivity.

Prior arithmetical knowledge

Students' arithmetical knowledge was measured before the intervention (T0). The measurement captured their knowledge of numbers, number relations, and the relation of addition and subtraction, as well as their competencies in addition (for sample tasks, see Supplementary file 6). The test consisted of 25 tasks and was scaled using a 1-PL logistic model for dichotomous data with the R-Package TAM (Robitzsch et al., 2018). The items had acceptable WMNSQ between 0.79 and 1.22. The EAP/PV- (0.82) and WLE-reliability (0.86) were satisfactory ($\sigma^2_{WLE} = 1.86$). The person-parameters were calculated using WLE ($M = -0.01$, $SD = 1.46$).

---

[2] The students did not use only the five idealized subtraction strategies taught in our treatment. Thus, our coding system also comprised shorter versions of the strategies, as well as longer versions with an additional solution step, combinations of two strategies, and misleading strategies.
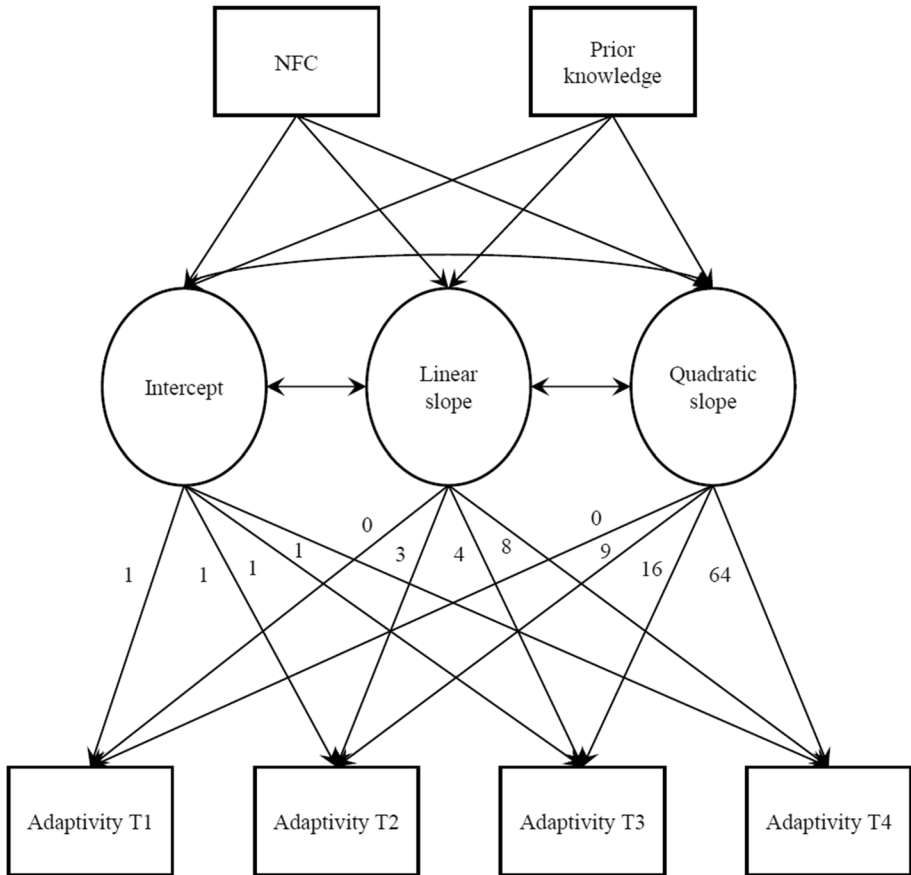
Fig. 3 Multiple-group LGCM (blocked vs interleaved) with NFC and prior knowledge as covariates

Need for cognition

We administered an NFC scale consisting of 10 items (Keller et al., 2019) before the intervention took place at T0 (Supplementary file 7). The items comprised self-reports about students' enjoyment of cognitive activities (e.g., "Thinking is fun for me") and about students' seeking of cognitively effortful activities (e.g., "I like to solve tricky problems"). The items were read aloud by the test leaders to counter possible reading difficulties of students. Items were answered on a 4-point Likert scale ranging from "Not true at all" to "Very true." Sample reliability was satisfactory ($\alpha=0.84$; $M=3.01$, $SD=0.58$).

Analyses

All research questions were addressed using multiple-group (interleaved vs blocked) latent growth curve models (LGCMs), which is a suitable statistical approach to evaluate experimental studies and, in particular, to analyze differences in development on the group and the individual level (Hesser, 2015). LGCMs provide the possibility to estimate mean growth trajectories, as well as between-person differences in development and potential

moderating effects of covariates on students' learning trajectories (Duncan & Duncan, 2004; Meredith & Tisak, 1990). The development is modelled by an intercept and slope factor(s). Given that the first time point is fixed to zero, the mean of the intercept represents the mean starting value of individuals, and the associated variance indicates whether there are interindividual differences regarding students' intercept. LGCMs can be extended by adding a slope factor, e.g., a linear slope, which describes the mean linear trend and interindividual differences in development. However, students do not linearly improve their competencies, and especially in intervention studies, their ability might decrease after the treatment when resuming their regular mathematics lessons. Therefore, we tested whether the inclusion of a quadratic slope factor is more appropriate to describe our data.

To detect the most appropriate LGCM, we compared the following multiple-group LGCM: intercept only; intercept and linear slope; intercept, linear slope, and quadratic slope. To take the different time intervals between the points of measurement into account, we specified the factor loadings of the slope factors under consideration of the weeks between the measures (T1: $\lambda = 0$, T2: $\lambda = 3$, T3: $\lambda = 4$, T4: $\lambda = 8$; Preacher et al., 2008). The factor loadings were squared for the quadratic slope factor.

To address our research questions as to whether students' prior arithmetical knowledge and NFC have a moderating effect on the effectiveness of interleaved/blocked practice on students' learning gains in using subtraction strategies adaptively, we extended our multiple-group LGCM by including students' prior arithmetical knowledge and NFC as covariates (Fig. 3). We used likelihood ratio tests to test for interaction effects of the covariates with teaching condition. All analyses were conducted with Mplus Version 8.5 using MLR-estimator and full information maximum likelihood (FIML) to deal with missing values.

Results

Descriptive statistics

Table 3 displays the latent means and standard deviations for students' adaptivity at the four points of measurement. A descriptive analysis of the development of the adaptive use of subtraction strategies reveals improvements from T1 to T3 and a decline from T3 to T4 in both conditions.

Table 3 Latent means and standard deviations in adaptivity separately for the interleaved and blocked condition

Condition		N	T1 M (SD)	T2 M (SD)	T3 M (SD)	T4 M (SD)
	Interleaved	119	−0.98 (1.28)	0.75 (1.22)	0.97 (1.19)	0.62 (1.28)
	Blocked	117	−0.93 (1.23)	−0.38 (0.98)	−0.11 (1.07)	−0.24 (1.01)
Total		236	−0.95 (1.26)	0.19 (1.25)	0.43 (1.26)	0.19 (1.23)

Development of the adaptive use of subtraction strategies

We compared different multiple-group LGCMs to detect the optimal growth function over the four points of measurement. Table 4 shows that a quadratic LGCM, which captures an increase as well as a decline in growth, describes students' learning trajectories best. Therefore, a quadratic LGCM was used for all subsequent analyses.

Table 5 shows the coefficients, standard errors, and p-values for the means, variances, and covariances of the intercept, slope, and quadratic slope. There was no significant difference between the two groups in the intercept, $\chi^2(1)=0.03$, $p=0.86$. The interleaved group exhibited a significant stronger linear growth in adaptivity (linear slope$=0.79$) than the blocked group (linear slope$=0.29$), $\chi^2(1)=35.72$, $p<0.001$. The means of the quadratic slope also differed significantly between the two conditions, $\chi^2(1)=33.49$, $p<0.001$, meaning that interleaved learning (quadratic slope$=-0.07$) was accompanied by a slightly stronger inhibition and decline of growth compared to blocked practice (quadratic slope$=-0.03$).

Table 4 Model comparison for multiple-group (interleaved vs blocked) LGCM

Model	χ^2	RMSEA	CFI	TLI	SRMR
Intercept only	$\chi^2(20)=187.60$, $p<0.001$	0.27	0.45	0.67	0.20
Intercept, slope	$\chi^2(15)=153.97$, $p<0.001$	0.28	0.54	0.64	0.17
Intercept, slope, quadratic	$\chi^2(8)=11.06$, $p=0.20$	0.06	0.99	0.99	0.04

Table 5 LGCM parameters for the interleaved and blocked condition in adaptivity

Parameter	Interleaved practice			Blocked practice		
	Coefficient	*SE*	*p*	Coefficient	*SE*	*p*
Mean						
Intercept	−0.98	0.12	<0.001	−0.95	0.12	<0.001
Slope	0.79	0.06	<0.001	0.29	0.05	<0.001
Quadratic	−0.07	0.01	<0.001	−0.03	0.01	<0.001
Variance						
Intercept	1.34	0.27	<0.001	1.18	0.26	<0.001
Slope	0.25	0.06	<0.001	0.21	0.05	<0.001
Quadratic	0.00	0.00	<0.001	0.00	0.00	<0.01
Covariance						
Intercept – Slope	−0.34	0.10	<0.01	−0.33	0.09	<0.001
Intercept – Quadratic	0.03	0.01	<0.001	0.03	0.01	<0.01
Slope – Quadratic	−0.02	0.01	<0.001	−0.02	0.01	<0.001

Notes: unstandardized coefficients are displayed

Effects of prior knowledge and NFC

Next, students' prior arithmetical knowledge (*research question 1*) and NFC (*research question 2*) were included in the multiple-group LGCM as predictor variables to investigate whether these variables interact with our learning conditions. The extended multiple-group LGCM had a satisfactory fit, $\chi^2(13)=15.14$, $p=0.30$, RMSEA$=0.04$, CFI$=1.00$, TLI$=0.99$, SRMR$=0.03$.

Students' prior arithmetical knowledge had a significant positive effect on the intercept in both the interleaved ($\beta=0.34$, 95% CI (0.16; 0.52), $SE=0.09$, $p<0.001$) and blocked condition ($\beta=0.21$, 95% CI (0.01; 0.41), $SE=0.10$, $p<0.05$), meaning that students with greater prior arithmetical knowledge solved the tasks more adaptively at T1. The difference in the parameters between the two groups was not significant ($\chi^2(1)=2.19$, $p=0.14$). Regarding linear growth, students' prior knowledge had a significant positive effect in the blocked condition ($\beta=0.26$, 95% CI (0.01; 0.50), $SE=0.12$, $p<0.05$) but not in the interleaved condition ($\beta=0.15$, 95% CI (-0.07; 0.37), $SE=0.11$, $p=0.17$). However, as the strongly overlapping confidence intervals show, the regression parameters did not differ significantly between the two groups, $\chi^2(1)=0.26$, $p=0.61$. Contrary to H1, with respect to linear growth no differential effect of prior knowledge for blocked and interleaved practice was found. Regarding the quadratic slope, prior arithmetical knowledge had no significant influence in both conditions (interleaved: $\beta=-0.11$, 95% CI (-0.34; 0.13), $SE=0.12$, $p=0.38$; blocked: $\beta=-0.23$, 95% CI (-0.51; 0.05), $SE=0.14$, $p=0.10$), and there was no significant difference between the two conditions regarding the effect on the quadratic slope, $\chi^2(1)=0.34$, $p=0.56$.

In both groups, students' NFC did not have a significant impact on intercept (interleaved: $\beta=-0.14$, 95% CI (-0.33; 0.04), $SE=0.10$, $p=0.13$; blocked: $\beta=0.10$, 95% CI (-0.13; 0.34), $SE=0.12$, $p=0.39$), and the two conditions did not differ regarding the effect of NFC on intercept, $\chi^2(1)=1.58$, $p=0.21$. Students' NFC affected linear growth only in the interleaved condition ($\beta=0.21$, 95% CI (0.01; 0.41), $SE=0.10$, $p<0.05$), but not in the blocked condition ($\beta=-0.09$, 95% CI (-0.32; 0.13), $SE=0.12$, $p=0.41$). This result indicates that the higher the students' NFC in the interleaved condition, the higher their learning gains, while it does not influence linear growth in the blocked condition. However, we have to reject H2 as the difference in the parameters failed to reach significance, $\chi^2(1)=3.51$, $p=0.06$. NFC had a significant influence on the quadratic slope in the interleaved condition ($\beta=-0.21$, 95% CI (-0.41; -0.00), $SE=0.11$, $p<0.05$), but not in the blocked one ($\beta=0.05$, 95% CI (-0.19; 0.30), $SE=0.13$, $p=0.68$). Nonetheless, the regression parameters did not differ significantly, $\chi^2(1)=2.96$, $p=0.08$.

Discussion

We investigated the role of students' prior arithmetical knowledge and NFC in the effectiveness of interleaved practice and blocked practice on students' adaptive use of subtraction strategies in third-grade mathematics. Regarding *research question 1*, our multiple-group LGCM has shown that students' prior arithmetical knowledge had a significantly positive influence on learning gains in the blocked condition (*research question 1*). Thus, a *Matthew effect* (Merton, 1968; Rigney, 2010; Simonsmeier et al., 2021) was found for blocked learning, meaning that higher prior arithmetical knowledge was accompanied by stronger linear growth in adaptivity. One reason for this result could be that the students in

our blocked condition were not explicitly prompted to compare the subtraction strategies (between-comparison). They were only prompted to decide whether tasks can be solved adaptively with a specific strategy (within-comparison). To select the most appropriate strategy for specific subtraction tasks, students need to compare multiple strategies to focus on and discern the central characteristics of the task (Kullberg et al., 2017; Marton & Pang, 2006). Therefore, it can be assumed that students with higher prior knowledge were more capable of engaging in these necessary comparison activities over a longer period and without our instructional support initiating between-comparisons, which we, on the contrary, provided in our interleaved condition. Although prior arithmetical knowledge had a significant impact on linear growth in the blocked but not in the interleaved condition, this difference is negligible as the parameters did not differ significantly. No significant differential effect of prior arithmetical knowledge on students' learning gains regarding adaptivity between interleaved and blocked practice was found. Therefore, H1, with which we expected that students' prior knowledge has a stronger impact on learning gains in blocked practice, has to be rejected.

Regarding *research question 2*, we investigated the role of students' NFC for interleaved and blocked practice. While students' NFC had a positive influence on the linear slope in the interleaved condition, meaning that students with a higher NFC had greater learning gains in the adaptive use of subtraction strategies, such an effect was not found in the blocked condition. Because interleaved practice combined with prompts to compare is more cognitively demanding than blocked learning, interleaved practice is more beneficial for students higher in NFC, i.e., with more pronounced enjoyment of and engagement in cognitively demanding activities. Our interleaved teaching approach probably matches these students' high NFC, leading to deeper information processing and thus higher learning gains (Evans et al., 2003). On the contrary, practicing different subtraction strategies successively, as in our blocked condition, does not urge students to engage in cognitively demanding comparison processes. However, the parameters between the two conditions did not differ significantly, although the confidence intervals were only slightly overlapping. Thus, we have to reject H2, with which we expected differential effects of students' NFC on interleaved and blocked practice.

The results related to the role of NFC in students' learning gains when learning subtraction strategies in an interleaved fashion raise the question of how students low in NFC can be supported to benefit more from interleaved practice. One possibility could be to gradually shift from a blocked to an interleaved practice of subtraction strategies, rather than implementing a pure form of interleaved practice. To be able to detect the differences between the tasks and thus decide which strategy is most adaptive, students need sufficient knowledge about when to use which strategy. This could be promoted by an initial blocked learning phase. Through an initial blocked phase, students get the chance to detect similarities between different subtraction tasks that can all be solved adaptively with the same subtraction strategy. The subsequent interleaved learning phase then offers students the chance to distinguish between subtraction tasks that are solved adaptively with different subtraction strategies. Hence, as students' proficiency increases, contextual interference and thus task difficulty increases as well (Nakata & Suzuki, 2019; Rohrer & Hartwig, 2020). Such an approach could reduce cognitive effort since students would start with a more pronounced prior knowledge regarding

the adaptive use of the different subtraction strategies in interleaved practice. Because research on increasing interleaved practice is inconsistent (e.g., Nakata & Suzuki, 2019; Pan et al., 2019; Sorensen & Woltz, 2016; Yan et al., 2017) and lacking for primary school mathematics, further research is required to investigate the assumed benefits of increasing interleaved practice for third-graders' adaptive use of subtraction strategies—especially for those low in NFC.

There are some limitations regarding our study that must be considered. First, due to our complex statistical models, statistical power was limited. Thus, smaller effects may have remained undiscovered, e.g., a possible differential effect of NFC on interleaved vs blocked practice, as suggested by the slightly overlapping confidence intervals of the two conditions regarding the effect of NFC on students' learning trajectories. Second, we combined interleaved practice with prompts to compare. Therefore, we cannot conclusively identify which of the instructional approaches is responsible for the impressive learning gains in our interleaved condition. Moreover, we cannot make a conclusive statement on whether prior arithmetical knowledge and NFC interact differently with interleaved/blocked practice and with comparison learning, since our intervention combined both instructional approaches. Third, our results indicate that interleaved practice combined with prompts to compare can promote students' adaptive use of subtraction strategies. However, further research needs to examine whether the positive effects of interleaved practice, including prompts to compare, and the role of prior knowledge and NFC, are transferable to other mathematical topics and other domains.

Declarations

Ethics approval and consent to participate This study was carried out in accordance with the recommendations of Declaration of Helsinki as well as the ethical guidelines of the German Psychologists Association (BDP) and the German Psychological Society (DGP). The protocol was approved by the Ethics Committee of the Faculty of Human Sciences (University of Kassel). All parents gave written informed consent in accordance with the Declaration of Helsinki.

Conflict of interest The authors declare no competing interests.

References

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

Baroody, A. J., & Dowker, A. (Eds.). (2003). *The development of arithmetic concepts and skills: Constructing adaptive expertise*. Erlbaum.

Birnbaum, M. S., Kornell, N., Bjork, E. L., & Bjork, R. A. (2013). Why interleaving enhances inductive learning: The roles of discrimination and retrieval. *Memory & Cognition, 41*(3), 392–402. https://doi.org/10.3758/s13421-012-0272-7

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). Worth Publishers.

Blöte, A. W., Klein, A. S., & Beishuizen, M. (2000). Mental computation and conceptual understanding. *Learning and Instruction, 10*(3), 221–247. https://doi.org/10.1016/S0959-4752(99)00028-6

Brunmair, M., & Richter, T. (2019). Similarity matters: A meta-analysis of interleaved learning and its moderators. *Psychological Bulletin, 145*(11), 1029–1052. https://doi.org/10.1037/bul0000209

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin, 119*(2), 197–253. https://doi.org/10.1037/0033-2909.119.2.197

Carvalho, P. F., & Goldstone, R. L. (2015). The benefits of interleaved and blocked study: Different tasks benefit from different schedules of study. *Psychonomic Bulletin & Review, 22*(1), 281–288. https://doi.org/10.3758/s13423-014-0676-4

Carvalho, P. F., & Goldstone, R. L. (2017). The sequence of study changes what information is attended to, encoded, and remembered during category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(11), 1699–1719. https://doi.org/10.1037/xlm0000406

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

Cook, G. L., & Odom, R. D. (1992). Perception of multidimensional stimuli: A differential-sensitivity account of cognitive processing and development. *Journal of Experimental Child Psychology, 54*(2), 213–249. https://doi.org/10.1016/0022-0965(92)90036-6

De Smedt, B., Torbeyns, J., Stassens, N., Ghesquière, P., & Verschaffel, L. (2010). Frequency, efficiency and flexibility of indirect addition in two learning environments. *Learning and Instruction, 20*(3), 205–215. https://doi.org/10.1016/j.learninstruc.2009.02.020

Duncan, T. E., & Duncan, S. C. (2004). An introduction to latent growth curve modeling. *Behavior Therapy, 35*(2), 333–363. https://doi.org/10.1016/S0005-7894(04)80042-X

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. H., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Durkin, K., Star, J. R., & Rittle-Johnson, B. (2017). Using comparison of multiple strategies in the mathematics classroom: Lessons learned and next steps. *The International Journal on Mathematics Education, 49*(4), 585–597. https://doi.org/10.1007/s11858-017-0853-9

Durkin, K., Rittle-Johnson, B., Star, J. R., & Loehr, A. (2023). Comparing and discussing multiple strategies: An approach to improving algebra instruction. *The Journal of Experimental Education, 91*(1), 1–19. https://doi.org/10.1080/00220973.2021.1903377

Evans, C. J., Kirby, J. R., & Fabrigar, L. R. (2003). Approaches to learning, need for cognition, and strategic flexibility among university students. *British Journal of Educational Psychology, 73*(4), 507–528. https://doi.org/10.1348/000709903322591217

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

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

Ginet, A., & Py, L. (2000). Le besoin de cognition: Une échelle française pour enfants et ses conséquences au plan sociocognitif [Need for cognition: A French scale for children and its socio-cognitive consequences]. *L'année Psychologique, 100*, 585–628. https://doi.org/10.3406/psy.2000.28665

Guo, J. P., Pang, M. F., Yang, L. Y., & Ding, Y. (2012). Learning from comparing multiple examples: On the dilemma of "similar" or "different." *Educational Psychology Review, 24*, 251–269. https://doi.org/10.1007/s10648-012-9192-0

Guo, J.-P., Yang, L.-Y., & Ding, Y. (2014). Effects of example variability and prior knowledge in how students learn to solve equations. *European Journal of Psychology of Education, 29*, 21–42. https://doi.org/10.1007/s10212-013-0185-2

Hartig, J., & Kühnbach, O. (2006). Schätzung von Veränderung mit „plausible values" in mehrdimensionalen Rasch-Modellen [Estimating change in multidimensional Rasch-models using „plausible values"]. In A. Ittel & H. Merkens (Eds.), *Veränderungsmessung und Längsschnittstudien in der empirischen Erziehungswissenschaft* (pp. 27–44). VS Verlag.

Heinze, A., Marschick, F., & Lipowsky, F. (2009a). Addition and subtraction of three-digit numbers: Adaptive strategy use and the influence of instruction in German third grade. *ZDM Mathematics Education, 41*(5), 591–604. https://doi.org/10.1007/s11858-009-0205-5

Heinze, A., Star, J. R., & Verschaffel, L. (2009b). Flexible and adaptive use of strategies and representations in mathematics education. *ZDM Mathematics Education, 41*(5), 535–540. https://doi.org/10.1007/s11858-009-0214-4

Heinze, A., Arend, J., Grüßing, M., & Lipowsky, F. (2018). Instructional approaches to foster third graders' adaptive use of strategies: An experimental study on the effects of two learning environments on multi-digit addition and subtraction. *Instructional Science, 46*(6), 869–891. https://doi.org/10.1007/s11251-018-9457-1

Heinze, A., Grüßing, M., Arend, J., & Lipowsky, F. (2020). Fostering children's adaptive use of mental arithmetic strategies: A comparison of two instructional approaches. *Journal of Mathematics Education, 13*(1), 18–34. https://doi.org/10.26711/007577152790052

Hesser, H. (2015). Modeling individual differences in randomized experiments using growth models: Recommendations for design, statistical analysis and reporting of results of internet interventions. *Internet Interventions, 2*(2), 110–120. https://doi.org/10.1016/j.invent.2015.02.003

Hickendorff, M. (2020). Fourth graders' adaptive strategy use in solving multidigit subtraction problems. *Learning and Instruction, 67*, 101311. https://doi.org/10.1007/s10212-017-0357-6

Hickendorff, M., McMullen, J., & Verschaffel, L. (2022). Mathematical flexibility: Theoretical, methodological, and educational considerations. *Journal of Numerical Cognition, 8*(3), 326–334. https://doi.org/10.5964/jnc.10085

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1–27). Erlbaum.

Keller, U., Strobel, A., Wollschläger, R., Greiff, S., Martin, R., Vainikainen, M.-P., & Preckel, F. (2019). A need for cognition scale for children and adolescents. *European Journal of Psychological Assessment, 35*(1), 137–149. https://doi.org/10.1027/1015-5759/a000370

Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding it up: Helping children learn mathematics*. National Academy Press.

Kullberg, A., Runesson Kempe, U., & Marton, F. (2017). What is made possible to learn when using the variation theory of learning in teaching mathematics? *ZDM Mathematics Education, 49*(4), 559–569. https://doi.org/10.1007/s11858-017-0858-4

Lemaire, P., & Siegler, R. S. (1995). Four aspects of strategic change: Contributions to children's learning of multiplication. *Journal of Experimental Psychology: General, 124*(1), 83–97. https://doi.org/10.1037/0096-3445.124.1.83

Loewenstein, J., Thompson, L., & Gentner, D. (1999). Analogical encoding facilitates knowledge transfer in negotiation. *Psychonomic Bulletin & Review, 6*(4), 586–597. https://doi.org/10.3758/BF03212967

Luong, C., Strobel, A., Wollschläger, R., Greiff, S., Vainikainen, M.-P., & Preckel, F. (2017). Need for cognition in children and adolescents: Behavioral correlates and relations to academic achievement and potential. *Learning and Individual Differences, 53*, 103–113. https://doi.org/10.1016/j.lindif.2016.10.019

Marton, F., & Booth, S. (1997). *Learning and awareness*. Lawrence Erlbaum Associates.

Marton, F., & Pang, M. F. (2006). On some necessary conditions of learning. *Journal of the Learning Sciences, 15*(2), 193–220. https://doi.org/10.1207/s15327809jls1502_2

McDaniel, M. A., & Butler, A. C. (2011). A contextual framework for understanding when difficulties are desirable. In A. S. Benjamin (Ed.), *Successful remembering and successful forgetting. A festschrift in honor of Robert A. Bjork* (pp. 175–198). Taylor & Francis.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55*(1), 107–122. https://doi.org/10.1007/BF02294746

Merton, R. K. (1968). The Matthew effect in science. The rewards and communication systems of science are considered. *Science, 159*, 56–63.

Nakata, T., & Suzuki, Y. (2019). Mixing grammar exercises facilitates long-term retention: Effects of blocking, interleaving, and increasing practice. *The Modern Language Journal, 103*(3), 629–647. https://doi.org/10.1111/modl.12581

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. The National Council of Teachers of Mathematics.

Nemeth, L., Werker, K., Arend, J., Vogel, S., & Lipowsky, F. (2019). Interleaved learning in elementary school mathematics: Effects on the flexible and adaptive use of subtraction strategies. *Frontiers in Psychology, 10*, 86. https://doi.org/10.3389/fpsyg.2019.00086

Nemeth, L., Werker, K., Arend, J., & Lipowsky, F. (2021). Fostering the acquisition of subtraction strategies with interleaved practice: An intervention study with German third graders. *Learning and Instruction, 71*, 101354. https://doi.org/10.1016/j.learninstruc.2020.101354

Pan, S. C., Tajran, J., Lovelett, J., Osuna, J., & Rickard, T. C. (2019). Does interleaved practice enhance foreign language learning? The effects of training schedule on Spanish verb conjugation skills. *Journal of Educational Psychology, 111*(7), 1172–1188. https://doi.org/10.1037/edu0000336

Preacher, K. J., Wichman, A. L., MacCallum, R. C., & Briggs, N. E. (2008). *Latent growth curve modeling*. Sage Publications.

Rau, M. A., Aleven, V., & Rummel, N. (2010). Blocked versus interleaved practice with multiple representations in an intelligent tutoring system for fractions. In V. Aleven, J. Kay, & J. Mostow (Eds.), *Intelligent tutoring systems* (pp. 413–422). Springer. https://doi.org/10.1007/978-3-642-13388-6_45

Rau, M. A., Aleven, V., Rummel, N., & Pardos, Z. (2014). How should intelligent tutoring systems sequence multiple graphical representations of fractions? A multi-methods study. *International Journal of Artificial Intelligence in Education, 24*, 125–161. https://doi.org/10.1007/s40593-013-0011-7

Richter, T., Berger, R., Ebersbach, M., Eitel, A., Endres, T., Borromeo Ferri, R., Hänze, M., Lachner, A., Leutner, D., Lipowsky, F., Nemeth, L., Renkl, A., Roelle, J., Rummer, R., Scheiter, K., Schweppe, J., von Aufschnaiter, C., & Vorholzer, A. (2022a). How to promote lasting learning in schools: Theoretical approaches and an agenda for research. *Zeitschrift Für Entwicklungspsychologie Und Pädagogische Psychologie, 54*(4), 135–141. https://doi.org/10.1026/0049-8637/a000258

Richter, T., Nemeth, L., Berger, R., Borromeo Ferri, R., Hänze, M., & Lipowsky, F. (2022b). Using interleaving to promote inductive learning in educational contexts: Promises and challenges. *Zeitschrift Für Entwicklungspsychologie Und Pädagogische Psychologie, 54*(4), 164–175. https://doi.org/10.1026/0049-8637/a000260

Rigney, D. (2010). *The Matthew effect: How advantage begets further advantage*. Columbia University Press.

Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology, 99*(3), 561–574. https://doi.org/10.1037/0022-0663.99.3.561

Rittle-Johnson, B., & Star, J. R. (2009). Compared with what? The effects of different comparisons on conceptual knowledge and procedural flexibility for equation solving. *Journal of Educational Psychology, 101*(3), 529–544. https://doi.org/10.1037/a0014224

Rittle-Johnson, B., Star, J. R., & Durkin, K. (2009). The importance of prior knowledge when comparing examples: Influences on conceptual and procedural knowledge of equation solving. *Journal of Educational Psychology, 101*(4), 836–852. https://doi.org/10.1037/a0016026

Rittle-Johnson, B., Star, J. R., & Durkin, K. (2012). Developing procedural flexibility: Are novices prepared to learn from comparing procedures? *British Journal of Educational Psychology, 82*(3), 436–455. https://doi.org/10.1111/j.2044-8279.2011.02037.x

Robitzsch, A., Kiefer, T., & Wu, M. (2018). TAM: Test analysis modules. R package version 3.0–21. Retrieved August 5, 2019, from https://CRAN.R-project.org/package=TAM

Rohrer, D., & Hartwig, M. K. (2020). Unanswered questions about spaced interleaved mathematics practice. *Journal of Applied Research in Memory and Cognition, 9*(4), 433–438. https://doi.org/10.1016/j.jarmac.2020.06.008

Rohrer, D., Dedrick, R. F., & Stershic, S. (2015). Interleaved practice improves mathematics learning. *Journal of Educational Psychology, 107*(3), 900–908. https://doi.org/10.1037/edu0000001

Rohrer, D., Dedrick, R. F., & Hartwig, M. K. (2020). The scarcity of interleaved practice in mathematics textbooks. *Educational Psychology Review, 32*(3), 873–883. https://doi.org/10.1007/s10648-020-09516-2

Selter, C. (2001). Addition and subtraction of three-digit numbers: German elementary children's success, methods and strategies. *Educational Studies in Mathematics, 47*(2), 145–173. https://doi.org/10.1023/A:1014521221809

Selter, C., Prediger, S., Nührenbörger, M., & Hußmann, S. (2012). Taking away and determining the difference – A longitudinal perspective on two models of subtraction and the inverse relation to addition. *Educational Studies in Mathematics, 79*(3), 389–408. https://doi.org/10.1007/s10649-011-9305-6

Simonsmeier, B. A., Flaig, M., Deiglmayr, A., Schalk, L., & Schneider, M. (2021). Domain-specific prior knowledge and learning: A meta-analysis. *Educational Psychologist, 57*(1), 31–54. https://doi.org/10.1080/00461520.2021.1939700

Sorensen, L. J., & Woltz, D. J. (2016). Blocking as a friend of induction in verbal category learning. *Memory & Cognition, 44*(7), 1000–1013. https://doi.org/10.3758/s13421-016-0615-x

Star, J. R. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education, 36*(5), 404–411. https://doi.org/10.2307/30034943

Star, J. R., & Rittle-Johnson, B. (2009). It pays to compare: An experimental study on computational estimation. *Journal of Experimental Child Psychology, 102*(4), 408–426. https://doi.org/10.1016/j.jecp.2008.11.004

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*(3), 185–223. https://doi.org/10.1207/s1532690xci1203_1

Taylor, K., & Rohrer, D. (2010). The effects of interleaved practice. *Applied Cognitive Psychology, 24*(6), 837–848. https://doi.org/10.1002/acp.1598

Thompson, L. A., & Markson, L. (1998). Developmental changes in the effect of dimensional salience on the discriminability of object relation. *Journal of Experimental Child Psychology, 70*(1), 1–25. https://doi.org/10.1006/jecp.1998.2445

Torbeyns, J., & Verschaffel, L. (2016). Mental computation or standard algorithm? Children's strategy choices on multi-digit subtractions. *European Journal of Psychology of Education, 31*(2), 99–116. https://doi.org/10.1007/s10212-015-0255-8

Torbeyns, J., De Smedt, B., Ghesquière, P., & Verschaffel, L. (2009). Acquisition and use of shortcut strategies by traditionally schooled children. *Educational Studies in Mathematics, 71*(1), 1–17. https://doi.org/10.1007/s10649-008-9155-z

Torbeyns, J., Hickendorff, M., & Verschaffel, L. (2017). The use of number-based versus digit-based strategies on multi-digit subtraction: 9–12-year-olds' strategy use profiles and task performance. *Learning and Individual Differences, 58*, 64–74. https://doi.org/10.1016/j.lindif.2017.07.004

Van der Auwera, S., De Smedt, B., Torbeyns, J., Verguts, G., & Verschaffel, L. (2023). Subtraction by addition in young multi-digit subtraction learners: A choice/no-choice study. *Journal of Experimental Child Psychology, 226*, 105544. https://doi.org/10.1016/j.jecp.2022.105544

Verschaffel, L., De Corte, E., Lamote, C., & Dhert, N. (1998). The acquisition and use of an adaptive strategy for estimating numerosity. *European Journal of Psychology of Education, 13*(3), 347–370. https://doi.org/10.1007/BF03172950

von Stumm, S., & Ackermann, P. L. (2013). Investment and intellect: A review and meta-analysis. *Psychological Bulletin, 139*, 841–869. https://doi.org/10.1037/a0030746

Yan, V. X., Soderstrom, N. C., Seneviratna, G. S., Bjork, E. L., & Bjork, R. A. (2017). How should exemplars be sequenced in inductive learning? Empirical evidence versus learners' opinions. *Journal of Experimental Psychology: Applied, 23*(4), 403–416. https://doi.org/10.1037/xap0000139

Zeileis, A., Strobl, C., Wickelmaier, F., Komboz, B., & Kopf, J. (2018). Package "psychotree". Recursive partitioning based on psychometric models. R package version 0.15–2. Retrieved August 5, 2019, from https://cran.r-project.org/web/packages/psychotree/psychotree.pdf

Ziegler, E., & Stern, E. (2014). Delayed benefits of learning elementary algebraic transformations through contrasted comparisons. *Learning and Instruction, 33*, 131–146. https://doi.org/10.1016/j.learninstruc.2014.04.006

Ziegler, E., & Stern, E. (2016). Consistent advantages of contrasted comparisons: Algebra learning under direct instruction. *Learning and Instruction, 41*, 41–51. https://doi.org/10.1016/j.learninstruc.2015.09.006

Ziegler, E., Edelsbrunner, P. A., & Stern, E. (2018). The relative merits of explicit and implicit learning of contrasted algebra principles. *Educational Psychology Review, 30*(2), 531–558. https://doi.org/10.1007/s10648-017-9424-4

Lea Nemeth. Department of Educational Science, University of Kassel, Nora-Platiel-Straße 1, 34109 Kassel, Germany. email: nemeth@uni-kassel.de

Current themes of research:

Interleaved practice. Learning by contrasting and comparing. Teacher education.

Most relevant publications in the field of Psychology of Education:

Götzfried, J., Nemeth, L., Bleck, V., & Lipowsky, F. (2022). Revision des Lerntypenmythos durch Konzeptwechselpodcasts im Lehramtsstudium [Dispelling the learning styles myth among pre-service teachers using conceptual change podcasts]. *Unterrichtswissenschaft, 50*(4), 639–659. https://doi.org/10.1007/s42010-022-00152-z

Nemeth L., Werker, K., Arend, J., & Lipowsky, F. (2021). Fostering the acquisition of subtraction strategies with interleaved practice: An intervention study with German third graders. *Learning and Instruction, 71*, 101354. https://doi.org/10.1016/j.learninstruc.2020.101354

Nemeth, L., Werker, K., Arend, J., Vogel, S., & Lipowsky, F. (2019). Interleaved learning in elementary school mathematics: Effects on the flexible and adaptive use of subtraction strategies. *Frontiers in Psychology, 10*, 86. https://doi.org/10.3389/fpsyg.2019.00086

Richter, T., Berger, R., Ebersbach, M., Eitel, A., Endres, T., Borromeo Ferri, R., Hänze, M., Lachner, A., Leutner, D., Lipowsky, F., Nemeth, L., Renkl, A., Roelle, J., Rummer R., Scheiter, K., Schweppe J., von Aufschnaiter, C., & Vorholzer, A. (2022). How to promote lasting learning in schools: Theoretical approaches and an agenda for research. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 54*(4), 135–141. https://doi.org/10.1026/0049-8637/a000258

Richter, T., Nemeth, L., Berger, R., Borromeo Ferri, R., Hänze, M., & Lipowsky, F. (2022). Using interleaving to promote inductive learning in educational contexts: Promises and challenges. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 54*(4), 164–175. https://doi.org/10.1026/0049-8637/a000260

Frank Lipowsky. Department of Educational Science, University of Kassel, Nora-Platiel-Straße 1, 34109 Kassel, Germany

Current themes of research:

Teacher professionalization. Teaching quality.

Most relevant publications in the field of Psychology of Education:

Bleck, V., & Lipowsky, F. (2022). Teachers' emotional exhaustion before and during the COVID-19 pandemic: Neither emotional exertion nor vacation feeling. *Frontiers in Psychology, 13*, 887494. https://doi.org/10.3389/fpsyg.2022.88749

Heinze, A., Arend, J., Grüßing, M., & Lipowsky, F. (2018). Instructional approaches to foster third graders' adaptive use of strategies—An experimental study on the effects of two learning environments on multi-digit addition and subtraction. *Instructional Science, 46*(6), 869–891. https://doi.org/10.1007/s11251-018-9457-1

Künsting, J., Neuber, V., & Lipowsky, F. (2016). Teacher self-efficacy as a long-term predictor of instructional quality in the classroom. *European Journal of Psychology of Education, 31*(3), 299–322. https://doi.org/10.1007/s10212-015-0272-7

Lipowsky, F. & Rzejak, D. (2015). Key features of effective professional development programmes for teachers. *Ricercazione. Six-monthly Journal on Learning. Research and Innovation in Education, 7*(2), 27–51.

Lipowsky, F., Rakoczy, K., Drollinger-Vetter, B., Klieme, E., Reusser, K., & Pauli, C. (2009). Quality of geometry instruction and its short-term impact on students' understanding of Pythagorean Theorem. *Learning and Instruction, 19*(6), 527–537. https://doi.org/10.1016/j.learninstruc.2008.11.001

Nägel, L., Bleck, V., & Lipowsky, F. (2023). "Research findings and daily teaching practice are worlds apart" – Predictors and consequences of scepticism toward the relevance of scientific content for teaching practice. *Teaching and Teacher Education, 102*, 103911. https://doi.org/10.1016/j.tate.2022.103911

Nemeth L., Werker, K., Arend, J., & Lipowsky, F. (2021). Fostering the acquisition of subtraction strategies with interleaved practice: An intervention study with German third graders. *Learning and Instruction, 71*, 101354. https://doi.org/10.1016/j.learninstruc.2020.101354

Nemeth, L., Werker, K., Arend, J., Vogel, S., & Lipowsky, F. (2019). Interleaved learning in elementary school mathematics: Effects on the flexible and adaptive use of subtraction strategies. *Frontiers in Psychology, 10*, 86. https://doi.org/10.3389/fpsyg.2019.00086