ORIGINALARBEIT/ORIGINAL ARTICLE

Children with Mathematical Learning Difficulties—How Do Their Working Memory Skills Differ from Typically Developing First Graders?

Kirsten Winkel [ORCID] · **Isabell Zipperle**

Abstract Children with mathematical learning difficulties differ from their peers not only in terms of their mathematics-specific competencies, but also in terms of other cross-curricular areas of competence. In many of these areas, they make more mistakes and need more time. Numerous studies show that they also make more mistakes in working memory tasks. However, there is little research on whether they also need more time in working memory tasks. The present study addresses this question. Methodologically, our study is aligned to peer studies that are included in a current meta-analysis. Our results from over 400 first graders reveal that children with mathematical learning difficulties do not only make more mistakes in working memory tasks, but also need significantly more time for these cognitive processes already. Our findings highlight how important it is to consider not only mathematics-specific competencies but also working memory skills and the time needed when diagnosing and supporting children in order to effectively implement individualized interventions and meet heterogeneous learning conditions in mathematics classroom in primary school.

Both authors contributed equally to this work.

✉ Kirsten Winkel · Isabell Zipperle
Johannes Gutenberg-Universität Mainz, Jakob-Welder-Weg 4, 55128 Mainz, Germany
E-Mail: kirsten.winkel@uni-mainz.de

⚫ Springer

Kinder mit Schwierigkeiten beim Mathematiklernen – Wie unterscheiden sich ihre Arbeitsgedächtnisfähigkeiten von denen ihrer Mitschülerinnen und Mitschüler in Klasse 1?

Zusammenfassung Kinder mit besonderen Schwierigkeiten beim Mathematiklernen unterscheiden sich von ihren Mitschülerinnen und Mitschülern nicht nur hinsichtlich ihrer mathematikspezifischen Kompetenzen, sondern ebenso hinsichtlich weiterer fächerübergreifend relevanter Kompetenzbereiche. In vielen dieser Bereiche machen sie mehr Fehler und benötigen mehr Zeit. Zahlreiche Studien belegen, dass sie auch bei Aufgaben zum Arbeitsgedächtnis mehr Fehler machen. Ob sie aber auch mehr Zeit bei Arbeitsgedächtnisaufgaben brauchen, ist bisher kaum erforscht. Die vorliegende Studie geht dieser Frage nach. Sie orientiert sich methodisch an den Studien, die in eine kürzliche Meta-Analyse einbezogen wurden. Unsere Ergebnisse von über 400 Erstklässlern zeigen, dass Kinder mit Schwierigkeiten beim Mathematiklernen nicht nur mehr Fehler bei Arbeitsgedächtnisaufgaben machen, sondern bereits für diese grundlegenden kognitiven Prozesse schon mehr Zeit benötigen. Unsere Ergebnisse unterstreichen, wie bedeutsam es ist, neben mathematikspezifischen Kompetenzen auch die Arbeitsgedächtnisfähigkeiten und die benötigte Zeit mit im Blick zu haben, um individuelle Förderpläne wirksam anzusetzen und den heterogenen Lernvoraussetzungen im Mathematikunterricht der Grundschule gerecht zu werden.

Schlüsselwörter Besondere Schwierigkeiten beim Mathematiklernen · Arbeitsgedächtnis · Benötigte Zeit · Arithmetik · Zahlensinn

1 Introduction

Specific difficulties in dealing with numbers are often hard to grasp for teachers and parents. Children have difficulties in understanding basic arithmetic concepts and use their fingers to solve basic calculations. However, many of them perform well in other school subjects. Such specific mathematical learning difficulties (MLD) are not a uniform phenomenon, but rather appear in many different forms (Von Aster and Lorenz 2013). If children are not encouraged consequently at an early age, this may have a negative impact on a child's future in and after school in many different areas. Standard classroom teaching is not enough for these children, and their difficulties remain even if they practice arithmetic tasks a lot. Even though the effects of untreated MLD are equally severe as those of other related learning impairments, there is much less research on MLD than on dyslexia or ADHD (Butterworth and Laurillard 2010). To gain a better understanding of MLD and to derive conclusions to support MLD children in school practice, research in mathematics education provides significant findings on diagnosis, prevention and support of MLD (e.g., Benz et al. 2017; Dornheim 2008; Fuchs et al. 2020; Graß and Krammer 2018; Häsel-Weide 2016; Schindler et al. 2020; Schipper 2002). Results emphasize the importance of effective mathematics teaching with natural differentiation for inclusive classrooms as well as of substantial learning environments for the development of

a sustainable understanding of numbers and numerical operations in primary school (Gaidoschik et al. 2021, p. 7 f.; Scherer et al. 2017).

At the same time, research in cognitive neuroscience and developmental psychology highlight the importance of more domain-general cognitive resources besides the importance of math-specific approaches and prior knowledge. The role of working memory (WM), in particular, has already been explicitly acknowledged in several theories of mathematical problem solving (Verschaffel et al. 2020, p. 8) and there is a wide consensus that mathematical ability does not only depend on skills in the mathematical domain, but that it is also related to other domain-general basic cognitive skills like intelligence, working memory, attention and language (Benz et al. 2017, p. 95; Schneider et al. 2021, p. 64 ff.). Working memory is generally responsible for holding information actively in mind while processing it and connecting new input to prior knowledge (Baddeley 1983). When learning mathematics or solving mathematical problems, several pieces of information like facts, tasks, questions, rules, patterns, numerical values, solution approaches or interim results must be held in mind at once, and linked together to understand mathematical contexts or to perform arithmetic or geometric operations (Passolunghi and Costa 2019). Likewise, WM capacity is needed when dealing with mathematical word problems (e.g., Fuchs et al. 2020). A child needs to make sense of the meaning of the text, memorize the key information, model the problem mathematically and keep all these pieces of information in mind to solve the problem. Since the capacity of human working memory—in contrast to the capacity of sensory memory and long-term memory—is clearly limited, it is also referred to as a "bottleneck for learning in classroom activities" (see Gathercole et al. 2006, p. 278). For children whose WM capacity is exhausted for a moment, their mathematical learning or problem solving process becomes harder, slower and more prone to error, even if the child does normally not have limitations in mathematical understanding (Passolunghi and Costa 2019).

Until recently, it was assumed that WM cannot be improved by training (e.g., Häsel-Weide 2016, p. 29). It is thus hardly surprising that WM has been given little attention in mathematics education. The effectiveness of WM training is still a subject of debate but increasing evidence over the past couple of years allows rejecting this assumption. Several studies reveal that WM can indeed be improved by training. A study in children with learning disabilities (Peijnenborgh et al. 2016) and two recent intervention studies with typically developing primary school children show that digital adaptive training can lead to increased visual-spatial WM, and this increased WM transfers to better achievement in arithmetic (Judd and Klingberg 2021) and in geometry over time (Berger et al. 2020). Likewise, regular mental arithmetic exercises with visual representations of numbers can cause an increase in visual-spatial WM capacity (Wang et al. 2019). It can be assumed that both WM and mathematical learning activity can influence one another (e.g., Benz et al. 2017, p. 95 ff.). This perspective on being able to promote WM skills in the classroom allows a reconsideration of WM within the field of mathematics education.

One general recommendation derived from research is that the MLD treatment should be adapted to individual children to increase the effectiveness of the intervention (e.g., DGKJP 2018a; Kaufmann and von Aster 2012). While claiming causality is difficult in the case of individual MLD, one can identify risk factors or predictors

for MLD (Schipper 2002). The evidence of WM being a strong and also malleable predictor of MLD (see Berger et al. 2020; Judd and Klingberg 2021) supports Schneider et al.'s argument (2021, p. 164) that WM should be included in the early diagnosis of MLD. In order to be able to customize MLD treatment after diagnosing WM difficulties, more detailed knowledge on interrelations between components of WM and specific mathematical skills is needed. The aim of the present paper is to contribute to this by using data from digital tests of accuracy and response times in all three WM components as well as tests on number and arithmetic processing in the first year of primary school.

Existing evidence on WM components as predictors for MLD is summarized in a meta-analysis on diagnosis and treatment of MLD (DGKJP 2018a[1]; Haberstroh and Schulte-Körne 2019). With reference to Ruwisch and Lorenz (2018, p. 43), the present paper takes this meta-analysis by Haberstroh and colleagues as a prompt to choose similar research methods and at the same time discuss the results from the perspective of mathematics education and school practice. However, before we go into more detail about our specific research questions derived from the aforementioned meta-analysis, we clarify how MLD is conceptualized in this study, document the current state of research on the association between WM and MLD and focus the specific research gap to be addressed in the present paper.

2 Background: Mathematical Learning Difficulties and Working Memory

2.1 Conceptualizing Mathematical Learning Difficulties

Mathematical ability has to be considered in a multidimensional way and requires a variety of skills such as knowledge of number facts, arithmetic procedures and concepts (Graß and Krammer 2018) as well as the ability to deal with multiple representations (Gagatsis and Shiakalli 2004). Specific difficulties in number and arithmetic processing are so widespread that they affect about 2 to 8% of all children (DGKJP 2018a). However, none of the existing terms related to these difficulties are used fully consistently nor are their definitions or inclusion criteria. Terms like 'dyscalculia' or 'Rechenstörung' (in German) are typically used in clinical and psychological approaches (Gaidoschik et al. 2021). In mathematics education, paraphrases like 'specific mathematical learning difficulties' are preferred to express that causes are not exclusively attributable to individual factors and that it is a multifaceted phenomenon with multiple causes (Gaidoschik et al. 2021, p. 4). Furthermore, in the latest version of the ICD-11 classification the term 'dyscalculia' is no longer used for young children (WHO 2022, code MB4B.5). The official clinical term has recently become "*developmental learning disorder with impairment in mathematics*" (WHO 2022, code 6A03.2). According to the World Health Organization, this develop-

---

[1] Results from this meta-analysis are published in parallel in German and in English language. The German publication by the DGKJP (2018a) and its appendix (DGKJP 2018b) is more detailed in some points. Main results are also documented in English language by Haberstroh and Schulte-Körne (2019).

mental learning disorder "*is characterized by significant and persistent difficulties in learning academic skills related to mathematics or arithmetic, such as number sense, memorization of number facts, accurate calculation, fluent calculation, and accurate mathematic reasoning*" (WHO 2022, code 6A03.2). Respectively, the diagnosis requires the exclusion of more severe intellectual or neurological disorders, vision or hearing impairments and a lack of language skills in the teaching language. Besides comparing mathematical performance to the child's age, the ICD-11 requires the application of an IQ discrepancy criterion by the condition of a minimum difference between mathematics achievement and the respective child's IQ. In contrast, scientists in mathematics education suggest not to use the IQ discrepancy criterion (Gaidoschik et al. 2021, p. 9) and the current discussion supports this suggestion with several scientific and practical arguments against the application of the IQ discrepancy criterion (e.g., Fischbach et al. 2013, p. 66; Maehler 2021, p. 219). Unaffected by this discussion, most studies do use a general IQ cut-off to separate between MLD and more severe disorders (e.g., of intellectual development). The purpose of this common approach is to derive particular conclusions for mathematics beyond comorbidities and low intellectual functioning. On the other hand, for an applied perspective in inclusive classrooms the cut-offs are to be considered as questionable since a child's need for mathematical support remains regardless of additional comorbidities.

In the domain of mathematics, several skills on number and arithmetic processing are affected in children with MLD (WHO 2022). Number sense as a measure for the development of the cardinal number concept is a strong predictor of MLD (Dornheim 2008; Jordan et al. 2010; Schneider et al. 2021, p. 74) and a recent eye tracking study identifies more counting strategies and longer response times in number sense tasks in children with MLD (Schindler et al. 2020). Basic arithmetic processing is one of the most fundamental skills to be learned in primary school mathematics and later mathematics instruction is constantly building on it.

2.2 State of Research on the Association Between Working Memory and Mathematics

Many studies in the field attest low WM to be a strong predictor of MLD (e.g., De Weerdt et al. 2013; Schuchardt et al. 2008; Szucs et al. 2013) and meta-analyses confirm these findings (Friso-van den Bos et al. 2013; Peng and Fuchs 2016). However, how exactly WM and mathematical ability are interrelated still lacks a conclusive answer (Viesel-Nordmeyer et al. 2020). Answering this question is challenging because cognitive profiles differ between children. The extent to which WM abilities are needed also depends on the mathematical problem, as well as the individual developmental stage and learning level of a child. The association between mathematical learning and WM evolves over time (Menon 2016). One reason for this is that the procedures and strategies used vary across different learning stages and mathematical topics, and that they require different levels of WM involvement. Geary (1993) points to a relationship with long-term memory: low WM capacities hinder children from correctly retrieving number facts from long-term memory and therefore trigger higher computational error rates. Difficulties in retrieving number

facts from long-term memory may result, however, in higher WM load. If WM fails or becomes overwhelmed, crucial information gets lost and learning becomes more difficult (Gathercole et al. 2016).

2.3 Working Memory and Its Components

The most referred-to WM model is developed by Baddeley and Hitch (1974). WM is a brain system "*needed for holding and manipulating information and for transferring it to the more permanent long-term memory system*" (Baddeley 1983, p. 74). In this model, WM is divided into three components: the central executive and its two subsystems, the phonological loop and the visual-spatial sketchpad. The central executive system forms the core of the model and serves as a supervisory system of limited capacity. It is responsible for the coordination of information provided by the phonological loop and the visual-spatial sketchpad: first, the phonological loop represents the verbal WM and covers the processing of verbal stimuli. Second, the visual-spatial sketchpad representing the visual WM is responsible for coping with visually presented information (Baddeley 1983). This multicomponent approach has been referred to in many studies which focus on the link between single components of the WM system and mathematical difficulties (for an overview see Dornheim 2008). The central executive system requires a high degree of cognitive flexibility in handing cues from different inputs. Tasks designed to measure the capacity of the central executive system are more complex than tasks designed to measure the capacity of the verbal or visual-spatial WM components. While evidence points to difficulties of the central executive system in children with MLD (see e.g., Schuchardt et al. 2008), findings for the verbal and visual subsystems are heterogeneous. Allen et al. (2020) claim that this may be the case because different components of WM are involved in different domains of mathematics. In general, the differences in WM capacity between children with difficulties in learning mathematics and their typically developing peers are larger for visual-spatial WM than for verbal WM. The studies included in the aforementioned meta-analysis by Haberstroh and Schulte-Körne (2019) show on average a significant medium or large effect size for the central executive system (Hedges' $g = 0.65$), a significant large mean effect size ($g = 0.84$) for the visual-spatial WM and a smaller, but still significant effect size ($g = 0.37$) for the verbal WM (DGKJP 2018a). Spatial abilities that are linked to visual-spatial WM explain roughly one fifth of the variance in arithmetical ability in the age group of primary school children (Graß and Krammer 2018). Considering the many visual representations and spatial patterns and structures in mathematics, it is plausible that visual-spatial WM capacity is essential for dealing with visually presented numbers and further spatially presented mathematical information (Toll et al. 2016). On the other hand, within the age group of 7- to 8-year-old children, Allen et al. (2020) found a higher correlation to verbal WM than to visual-spatial WM.

2.4 Accuracy and Response Times Measures

When measuring differences between children with MLD and their typically developing peers, there are two relevant measures in quantitative research in the field: Accuracy (or error rates) and response times. The meta-analysis by Haberstroh and colleagues (DGKJP 2018a; Haberstroh and Schulte-Körne 2019) lists these two outcome dimensions separately in a large table for various test measures. For example, in tests on numerical processing, basic arithmetic operations or word problems, children with MLD make significantly more mistakes and they also need more time to solve the tasks. Likewise, children with MLD make more mistakes in all tests on WM. According to Ranger and Kuhn (2012), accounting for response times is increasingly important. Gordon et al. (2020) provide evidence for 7- to 8-year-old children that response times in working memory tests are predictive for performance. Their results suggest that deficiencies in WM not only channel through lower accuracy measures (or higher error rates, respectively) but also through longer response times in the same tests. Due to the limited capacity of human WM, the major challenge in WM tasks is to temporarily keep several pieces of information in mind for a short period of time and to relate them to each other. A longer time period in turn increases the probability of losing task-critical information and thus making additional mistakes (Cornoldi and Giofrè 2014). Hence, fast processing in WM is crucial to keep information active in WM (Gordon et al. 2020). For arithmetic learning, this means that if numbers, interim results and operations are processed faster, the probability of losing information during the calculation process is lower. The same task could be more challenging for a child who needs more time for the answer because this child has to hold the same task-critical information (e.g., numbers) actively in mind for a longer time.

Response time in WM tasks needs to be distinguished from response time in processing speed tasks. Processing speed is a related but distinct concept that addresses mental processing at a very low cognitive level (Fry and Hale 2000). Processing speed measures the time a person needs from being presented a stimulus over accessing this incoming information to responding to this information (ibid.). For example, processing speed in this basic sense is the time between the moment when three cards with single digit numbers are presented and the moment when the child taps the card with the largest number (when assuming that the child has developed an ordinary understanding of numbers already). Importantly, processing speed measures comprise only very basic tasks between stimulus perception and response. Hence, processing speed cannot be a measure for number fact retrieval because the task can be considered simple enough only when it is assumed that all children have memorized the numerical facts. Processing speed is thus neither a measure of the ability to recall number facts from long-term memory nor is it relevant for solving mathematical problems.

Despite this importance of the time dimension in information processing in WM, research on MLD predictors still lacks studies with this focus (DGKJP 2018a, p. 15). The only study included in the meta-analysis by the DGKJP was conducted by De Weerdt et al. (2013) and reported response times for their sample of 112 children, but they could not detect any differences between children with MLD and controls.

To our knowledge, there are no recent studies focusing on response times in WM in children with MLD either.

3 Objective and Research Questions

Our study contributes to the strand of literature that investigates the relationship between MLD and WM. It aims to provide new evidence on the cognitive correlates of MLD. More specifically, we focus on accuracy and response time differences between children with and without MLD with respect to the central executive system and verbal and visual-spatial WM. Based on the findings of Gordon et al. (2020), who emphasize the role of response times in WM as a measure for higher order cognitive processing speed, we hypothesize that children with MLD do not only differ regarding accuracy but also regarding response times in WM tasks. Thus, we expect both accuracy and the time a child needs for higher order cognitive processing to be related to the ability to process numbers and operations.

This leads us to the following two research questions:

1. How do children with mathematical learning difficulties differ from typically de-veloping peers regarding their accuracy in working memory tests? Does our data replicate current evidence?
2. How do children with mathematical learning difficulties differ from typically de-veloping peers regarding their response times in working memory tests?

4 Methodology

4.1 Data

Our sample contains over five hundred first grade students ($N = 572$) from 12 primary schools in Mainz/Germany. All tests (except IQ) were specifically developed and programmed for our main study (Berger et al. 2020), which focused on causal effects of a randomized WM intervention. The present paper only uses correlational data from the first testing period before the intervention started. Data collection was conducted in the schools in groups of five children, supervised by interviewers experienced in standardized testing procedures with children.

In order to obtain precise time and accuracy data from children who cannot yet read or write fluently, we conducted digital tests and used large touchscreens and headphones with fully programmed instructions and practice trials for testing the children. Children entered their responses by tapping on the screen with their fingers. All inputs and precise response times were logged digitally. The children were tested in spring after they had been in school for half a year; mean age at test was 7.0 years. Children completed several tests on academic and cognitive skills. Full data for all relevant variables was available for 543 children. For this study, we use data from tests on number and arithmetic processing and WM as well as tests on reading comprehension and fluid IQ. Interviewers were blind to classroom behavior

and teachers were blind to testing. The class teachers filled out a questionnaire for each child. However, in the current study, we mainly use the digital test data from the children.
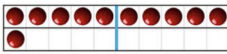
4.2 Measures

4.2.1 Measures for Mathematical Learning Difficulties

Conceptualizing MLD as difficulties in number and arithmetic processing and given, that these skills are multidimensional, any test can only capture a subset of the broad range of mathematics. Standardized tests for MLD differ in mathematical constructs and tasks used and in the way of implementation and evaluation (see Schneider et al. 2021). Given the limited reading and writing skills in our sample's age group, the implementation as a group test, and limited testing time, we did not include tests on mathematical reasoning or word problem solving in our study. We operationalize MLD as follows: we focus on number sense related to the cardinal number concept and on basic arithmetic processing in addition and subtraction up to 20. Hence, when we talk about MLD in our study, we are referring to learning difficulties in number and arithmetic processing. Our digital tests were provided in such a way that tasks were presented on the screen and via headphones and children were asked to enter resulting numbers into an input matrix on the touchscreen. The digits on this matrix were arranged in a way that they could not be used as a visual counting aid (see Fig. 1).

According to the curriculum for the second half year of grade 1, some easy items were in the number range up to 10 and all items were in the number range up to 20. The difficulty level varies across items. We used a number sense task as well as an auditory arithmetic task with 10 items each and a written arithmetic task on addition and subtraction with 11 items to assess arithmetic skills. Figure 1 demonstrates examples of the three mathematical subtasks.

The number sense task aims to measure number processing skills based on the cardinal number concept. The task demands subitizing skills when there are just a few balls and groupitizing skills when for example $5+3$ balls are displayed for a short time (Dornheim 2008, p. 257; Schindler et al. 2020). The balls were presented for 1.7 seconds on the screen. This display duration is typically too short to

Mathematical construct	Mathematical subtasks	Example tasks	Presentation of stimulus	Input Matrix
Cardinal number concept	Number sense task		A 20-frame with a number of balls is presented for 1.7 sec on the screen	
Basic arithmetic processing (addition and subtraction)	Auditory arithmetic task	"How much is 10 minus 4?" or "How much is 13 plus 2?"	Blank screen with auditory stimulus via headphones	(permanently visible for touchscreen input)
	Written arithmetic task	18-17=___ or 1+5+4=___	Task presented on the screen	

Fig. 1 Examples for digital mathematical subtasks for the assessment of number and arithmetic processing skills

count the balls. Hence, children had to internalize and structure the visual number representation and operate with the numbers in mind in order to determine the number of balls. We used this number sense task because it is known as a strong math-specific predictor of MLD (Schneider et al. 2021, p. 74).

In line with other standardized tests for the assessment of basic arithmetic skills (for an overview see Schneider et al. 2021, p. 171 ff.), we collected data on basic arithmetic processing. To this end, we used two tasks on addition and subtraction, an auditory one and a written one. The auditory task included typical mental arithmetic tasks with two numbers to add and subtract. Each test item in the auditory task was presented only once and required children to memorize and process verbal numerical information. In the written arithmetic task, more difficult items requiring addition and subtraction of several numbers were included. Most items in this task went beyond pure retrieval of number facts. Adapted to the fact that children with MLD often use counting strategies (e.g., Häsel-Weide 2016), we selected items involving obvious disadvantages when using counting strategies and obvious advantages for flexible strategy use. For example, if a child solves the task "$1+5+4=$" by counting without using the commutative law of addition, the child might begin at 1, count 5 steps forward (with fingers) and then count 4 steps forward. This takes time and includes the risk of miscounting. Another child with skills on flexible strategy use might combine the first and the last number to the number fact "$1+4=5$" by applying the commutative law intuitively. Then the child can recall the remaining task "$5+5=10$" as a number fact from long-term memory, too. Hence, the solution from the second child is expected to be much faster and less error-prone than the counting strategy from the first child.
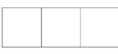
All three mathematical subtasks correlate significantly with each other ($p < 0.001$). This indicates a common underlying construct and allows for combining into a composite score for mathematical skills. As suggested by Rousselle and Noël (2007), we generated our composite score by adding up the scores from all three subtasks.

Additionally, we use teacher assessments of mathematical abilities for each child to validate our math test measures. Teachers were asked to rate children's overall ability in math on a scale from 1 to 7. We find a highly significant ($p < 0.001$) correlation coefficient of around $r = 0.6$. Since teachers were blind to tests and interviewers were blind to teaching, we consider this correlation to be a good indicator for the robustness of our math test scores.

4.2.2 Measures for Working Memory

The selection of tests was motivated by the intention to cover a broad spectrum of WM skills, including the central executive, the phonological loop and visual-spatial sketchpad as well as different degrees of complexity (two complex tasks and one simple) while sticking to the limit of a total of three different WM tasks. The WM tasks are described in Fig. 2.

We used a complex span task on memorizing auditorily presented words as a measure for the central executive (comp. DGKJP 2018a, p. 67). We used a simple span task on memorizing a series of digits as a measure for a math-related WM demand in the phonological loop. To measure the capacity of the visual-spatial sketchpad,

WM subcategory (test)	Example tasks	Description	Response screen
Central executive (verbal complex span task)	Auditory stimulus "*pig*" via headphones + button to press for animal vs. no animal, "*cake*" + button to press, "*elephant*" + button to press	After each auditorily presented word, the child was asked to push a button "Animal" or a button "No Animal". After the series of words, the child was asked to tap the corresponding pictures in the same order they heard the words.	
Phonological loop (verbal simple span task)	Auditory stimuli "*seven*", "*three*", "*eight*", "*one*" via headphones	After the auditorily presented series of numbers (with a pause of 2 sec between them), the child was asked to tap all numbers in the same order they heard them before.	1 2 3 4 5 6 7 8 9
Visual-spatial sketchpad (visual-spatial complex span task)	Screen 1 Screen 2 Screen 3	Three shapes were displayed at a time, two of which were the same. The child was asked to tap the "odd" shape, which was easy to distinguish at first glance. After the series of shapes, the child was asked to tap the positions of the "odd" shapes in the same order (e.g., middle - right - middle) in an empty grid (= WM demand).	

Fig. 2 Examples for the three working memory subtests

we used a complex span task in which children had to memorize the positions of shapes in a row. In all three WM tasks, the difficulty level was increased by varying the number of stimuli within an item stepwise from 2 to 7 in the two complex span tasks and from 2 to 9 in the simple span task.

4.2.3 Processing Speed

In line with processing speed tests used in related literature (e.g., Bull and Johnston 1997), our task did neither challenge WM nor intelligence nor mathematics skills or understanding (see Fry and Hale 2000). Our processing speed measure was taken as a separate side measure within the visual-spatial span task which was not included in WM measures. It is based on the child's selection of the 'odd' shape. We measure the time children needed from getting presented the three shapes until they chose the 'odd' shape. Although children were not explicitly asked to respond quickly, a comparison of average response times per item and of response time variability between the sample in Bull and Johnston (1997, p. 12) and our sample allows us to assume that children recognized the 'odd' shape at first glance and tapped on it immediately.

4.2.4 Construction of Accuracy and Response Time Measures

In all tests, we chose equivalent measures for the scale "Accuracy" ("number of solved items, or number of errors") and for the scale "Response Time" ("time needed for solving") as in the meta-analysis (see Haberstroh and Schulte-Körne 2019, p. 109). Our accuracy measures were the number of correctly solved items in a test. We z-standardized these measures over the whole representative sample with a mean of 0 and a standard deviation of 1 to account for better comparability across tests. Our measures for the time scale were based on the precise response times in the WM tasks. For each child, we calculate average response times for each item in each of our three WM tasks. As an additional proof of robustness for our results, we

evaluate our response time measures in two different ways: our main response time measure is based on all items of the respective test. We compute a second response time measure based only on the correctly solved items to control for a student's outcome orientation (see Landerl et al. 2004; Rousselle and Noël 2007). To make scores of tests with different content, times, and levels of difficulty comparable, we z-standardize our response time measures on the level of each single item.

4.3 Characteristics of MLD and Control Group

We follow the most common procedure to investigate differences between children with MLD and typically developing children but we do not apply the IQ discrepancy criterion in the present paper for the abovementioned arguments by Gaidoschik et al. (2021, p. 9).

In the pertinent literature, it has become an established procedure to define cut-off criteria, split up the sample into two groups, one group of children with MLD and one group of typically achieving children, and compare statistical differences between them. More specifically, our cut-off criteria exclude children with a fluid IQ below the 9th percentile (comparable to an IQ < 80), children with reading skills below the 16th percentile and children who are reported as having severe problems with the language of instruction by their teacher. Fluid intelligence was measured based on a subset of 17 items of Raven's Matrices IQ test scores (Bullheller and Häcker 2002). Reading comprehension was measured on the level of sentences with an age-appropriate test with 10 items of varying difficulty. Children had to choose the missing word in a sentence with a gap from a list of four alternatives (e.g., "Leo is at the [_____]." with four words to choose from: [Mum], [lake], [hat], [name]).

After applying the cut-offs, our sample consists of children with typical intelligence and reading skills who can follow the teaching. The cut-offs are typically applied in order to accentuate difficulties in the area of mathematics versus other difficulties. In addition to the procedure with these typically applied exclusions, we document parallel analyses on our full sample without excluding any children. By doing so, we aim to generalize our results for an applied perspective in inclusive classrooms in primary schools.

In all our analyses, children are assigned to the group of children with MLD if their mean score in tests on number and arithmetic processing is below the 25th percentile. In doing so, we follow the majority of peer studies, although this percentile exceeds estimated prevalence rates (for a discussion see Murphy et al. 2007). Our full sample—analyzed in the online supplement—includes 543 children with full sets of available data. After applying the cut-off criteria, we have a final sample of 409 children for the analyses in the main paper. 57 children with mathematical learning difficulties are assigned to the MLD group and the remaining 352 children to the control group. Table 1 provides summary statistics for these two groups.

4.4 Statistical Procedure

Our statistical procedure also follows the procedure of most peer studies (comp. DGKJP 2018b). First, we conduct correlation analyses and simple univariate t-tests

Table 1 Descriptive statistics by group

	N	Mean	SD	N	Mean	SD
	Control group			MLD group		
Accuracy central executive	352	5.18	1.85	57	4.12	2.21
Accuracy verbal WM	352	6.04	2.04	57	4.72	1.89
Accuracy visual-spatial WM	352	4.65	1.85	57	3.40	1.67
Response time central executive	352	1060	323	57	1105	378
Response time verbal WM	352	695	203	57	772	216
Response time visual-spatial WM	352	459	189	57	508	198
Response time processing speed	352	741	143	57	774	152
Math composite score	352	21.4	4.03	57	10.5	2.24
Number sense score	352	7.87	1.51	57	5.00	1.73
Auditory arithmetic score	352	7.6	1.93	57	3.18	1.31
Written arithmetic score	352	5.97	2.23	57	2.28	1.49
Reading score	352	8.09	1.69	57	7.21	1.82
IQ score	352	106	11.6	57	98.5	11.0
Age (in years)	352	7.01	0.36	57	7.03	0.36
Female (= 1)	352	0.53	–	57	0.58	–

Accuracy measured in number of correctly solved items (with a maximum of 12 in the verbal WM task, 11 in the other two WM tasks and in the written arithmetic task, 10 in the other two math tasks and in the reading task); response time is averaged per item and measured in milliseconds. The number of outliers (values beyond 3 SD deviations from the mean) amounts to less than 1%

to investigate the relationship between our mathematical and WM measures. Further t-tests are performed to check whether the groups differ from the mean of the population. Second, we run regression analyses of our accuracy and response time measures in single WM components on a group indicator and a list of covariates including demographic variables (age and gender) and fluid intelligence. Furthermore, to allow for a more differentiated view at the level of mathematical subtasks, we include continuous measures of mathematical ability in a further specification of our analysis.

Besides these main analyses and the parallel analysis with the full sample, we run different robustness checks. In a first robustness check, we include the response time measure in our accuracy analysis and vice versa to disentangle them. In a second robustness check, we run our regression using an alternative response time measure that is calculated on only correctly solved items. Lastly, we run regressions including teacher and school fixed effects to control for systematic differences between schools.

5 Results

Correlation analysis shows positive associations between both accuracy and response time measure in mathematical tests and in WM tests. Moreover, our t-tests indicate that the deviation of our MLD from our estimated population mean is positively on accuracy measures and negatively on response time measures (see online supplement

Table 2 Differences in the accuracy (number of solved problems) in WM tasks between children with and without mathematical learning difficulties

	Accuracy central executive		Accuracy verbal WM (phonological loop)		Accuracy visual-spatial WM (visual-spatial sketchpad)	
	β	p	β	p	β	p
MLD vs. control	–0.388*	(0.019)	–0.468***	(0.000)	–0.481***	(0.000)
Constant	0.965	(0.368)	2.218*	(0.018)	1.368	(0.088)
Controls	Yes		Yes		Yes	
Observations	409		409		409	
R-squared	0.072		0.106		0.114	

MLD vs. control = group dummy indicating the difference between MLD group and control group; The variance in the dependent variables explained by the group dummy and controls is limited. This is obvious regarding the binary structure of our independent variable and the low number of explanatory variables given the complexity of dependent measures. SE clustered on class level; *p*-values in parentheses

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

Table OS.1 to OS.3). Based on these basic findings, we conduct regression analyses to test whether group differences are significant. Our main interest variable is the group dummy 'MLD vs. control'. This dummy variable serves as an indicator for the difference between the group of children with mathematical difficulties and typically achieving control children. The coefficient can be interpreted as the accuracy or response time difference of a child being in the MLD group compared to a child being in the control group. Our robustness checks in the online supplement overall provide evidence that our results are robust to changes in model specification (see online supplement Table OS.8 to OS.16).

5.1 Accuracy of Working Memory Tasks

Our regression analysis links the accuracy measures (number of correctly solved items) of each of our three WM tests to their affiliation with the MLD group (see Table 2). Children in the MLD group on average scored significantly lower (0.39 to 0.48 SDs) than children in the control group in all three WM subtasks while keeping control variables constant in the linear regressions.

Our results from the parallel analyses on our full sample without excluding any children are documented in the online supplement (Table OS.4). The results remain almost the same—regardless of whether we apply cut-off criteria or not. This allows to generalize our results to the full sample.

5.2 Response Times in Working Memory Tasks

In order to analyze the time needed in WM tasks, we regress response time measures in the mathematical and WM tests on our group dummy variable and on controls. Apart from using time measures as dependent variables, regressions are identical to the regressions we run for exploring coherence on the accuracy scale.

Table 3 presents results for the regressions of mean response times on the group dummy and on controls. The coefficients are smaller than those from the accuracy

Table 3 Differences in response time in WM tasks between children with and without MLD

	Response times central executive		Response times verbal WM (phonological loop)		Response times visual-spatial WM (visual-spatial sketchpad)		Response times processing speed	
	β	p	β	p	β	p	β	p
MLD vs. control	0.119	(0.110)	0.223**	(0.001)	0.204*	(0.028)	0.145	(0.111)
Constant	0.946*	(0.014)	0.851*	(0.030)	0.991	(0.097)	1.096	(0.094)
Controls	Yes		Yes		Yes		Yes	
Observations	409		409		409		409	
R-squared	0.026		0.044		0.023		0.034	

MLD vs. control = group dummy indicating the difference between MLD group and control group; time measures are constructed based on all items; SE clustered on class level; *p*-values in parentheses
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

measures, but we find significant differences in response times in the tasks measuring two components of WM, the verbal WM task (0.22 SD, $p < 0.001$) and the visual-spatial WM task (0.2 SD, $p < 0.05$). Translated into actual response times, the mean group differences amount to 77 milliseconds in the verbal WM task and 49 milliseconds in the visual-spatial WM task respectively. The coefficients in the central executive and in the processing speed model are smaller and insignificant.

5.3 Detailed Results for Mathematical Subtasks

To provide a differentiated view on the mathematical tasks, we rerun our regressions with a continuous composite measure of mathematical ability instead of the group dummy as well as a continuous measure of the mathematical subtasks. We find that

Table 4 Differences of accuracy and response times in working memory tasks, using the math composite score and mathematical subtasks as independent variables

	Accuracy central executive	Accuracy phono-logical loop	Accuracy visual-spatial sketchpad	Response time central executive	Response time phono-logical loop	Response time vi-sual-spatial sketchpad
Math composite score (all 3 sub-tasks)	0.240***	0.319***	0.295***	–0.064*	–0.157***	–0.126**
Number sense task	0.118	0.165**	0.182***	–0.057	–0.112***	–0.089*
Auditory arith-metic task	0.249***	0.259***	0.166**	–0.066*	–0.132***	–0.105***
Written arith-metic task	0.160**	0.276***	0.302***	–0.023	–0.108**	–0.087*

The table shows coefficients from various separate regressions. In all regressions we controlled for gender, IQ and age. SE clustered on class level
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

Table 5 Comparison of effect sizes of belonging to the MLD group for WM components in the meta-analysis (DGKJP 2018a, p. 17) and in our present study

Main category	Subcategory	Effect size accuracy		Effect size response time	
		Meta-analysis	Present study	Meta-analysis	Present study
Working Memory	Central executive	0.65	0.56	–	0.18
	Phonological loop	0.37	0.65	–	0.42
	Visual-spatial sketchpad	0.84	0.68	–	0.31

The reported effect size in each category resembles the absolute value of the Hedges' g estimator

accuracy and response times in the verbal and visual-spatial WM are significantly associated with all three mathematical subtasks (see Table 4).

5.4 Effect Sizes

In Table 5, we translate our results into effect sizes to compare with the effect sizes reported in the meta-analysis which was published by the DGKJP in 2018 in German (DGKJP 2018a) and 2019 by Haberstroh & Schulte-Körne in English language (Haberstroh and Schulte-Körne 2019).

For the first dimension of accuracy, the effect sizes we obtain for the central executive and the visual-spatial sketchpad are similar to the effect sizes reported in the meta-analysis (ibid.). In the WM subcategory phonological loop, our estimated effect size is higher than the effect size reported by the meta-analysis (ibid.). For the dimension of response times, the meta-analysis stated missing evidence related to WM (DGKJP 2018a, p. 15). Our results indeed show a difference between children with and without MLD also regarding response times in WM tasks. Compared with the accuracy dimension, the effect sizes for response times are generally smaller. However, in the verbal and the visual WM components, small to medium effect sizes can be found, although we are talking about times of mostly less than one second per item with group differences of 4 to 11% (comp. Table 1).

6 Discussion

The present study aims to shed light on the association between MLD and WM skills—beyond fluid intelligence—in first grade children. Our research question addresses differences in accuracy and response times between children with and without MLD.

6.1 Differences in Accuracy of Working Memory Tasks

Concerning our first research question on differences in the accuracy of WM tasks, our study reveals significant differences in all three WM components, the central executive, the verbal and the visual-spatial WM, whereas some studies (e.g., Maehler and Schuchardt 2016; Schuchardt et al. 2008) report significant differences in single

WM components only. Our estimated effect size for the phonological loop is much larger than the mean effects size in the meta-analysis by Haberstroh and Schulte-Körne (2019). A potential explanation for this finding might be the young age of children in our sample. Children in the majority of studies included in the meta-analysis were older (DGKJP 2018b). Menon (2016) argues that verbal WM plays a prominent role particularly in the early stages of mathematical learning when children learn to verbally relate numbers and quantities to each other. However, these results must be viewed with caution due to the math-related content of our digit span task and the fact that we do not have a parallel measure from a letter span task. Hence, it can be assumed that the detected large effect size could be an overestimation. Concerning the central executive, our results confirm the finding that children with MLD make more mistakes in these WM tasks (see e.g., De Weerdt et al. 2013; Maehler and Schuchardt 2016; Schuchardt et al. 2008). For visual-spatial WM, this statement applies as well (see Table 5). This strong effect size is in line with a substantial number of studies indicating that MLD is particularly associated with problems in the visual-spatial WM component (e.g., Graß and Krammer 2018; Maehler and Schuchardt 2016; Szucs et al. 2013). In summary, our results confirm the importance of low WM skills as a strong predictor of MLD.

6.2 Differences in Response Times in Working Memory Tasks

Based on recent findings on the crucial need for fast processing in WM to keep information actively in mind (Gordon et al. 2020), we focus on response times as a measure for higher order cognitive processing speed. Our main contribution to the current literature is the following finding: limitations in WM in children with MLD might not only channel through higher error rates but also through longer response times in WM. The effects sizes in the time dimension are smaller than the effect sizes in the accuracy dimension but children with MLD needed significantly more time for their responses in two out of three WM tasks compared to their typically achieving peers. The use of touchscreens has made it possible to detect these small differences in response times.

Since we are talking about differences of around 50 milliseconds, we assume that we might not have been able to detect these small differences without using touchscreen as an input device, because prior experience of children with using computer mice might have caused disturbing noise in the time data (comp. De Weerdt et al. 2013).

Our results for response time differences reveal an effect size of 0.42 for verbal WM and of 0.31 for visual-spatial WM. The coefficients for the central executive in our main specification point in the same direction but are insignificant. Hence, we take a closer look at our WM tasks for verbal and visual-spatial WM. Children with MLD needed more time than their typically achieving classmates to indicate the spatial positions of objects although they were not significantly slower in distinguishing these objects before. The control measure for processing time is insignificant. With respect to the close relationship between WM and processing time (Cornoldi and Giofrè 2014), this finding suggests that children with MLD are not generally slower in cognitive processing, but they are slower when it comes to higher order cognitive

processing where several pieces of information have to be processed at a time. Concerning mathematical learning, it is plausible that this kind of higher order cognitive processing speed might be related to e.g., multistep arithmetic operations whereas our more basic control measure for processing speed might be more related to retrieval of e.g., basic number facts. Unfortunately, we did not collect systematic data on number fact retrieval to prove this hypothesis. Accordingly, we can only speculate that children with MLD are slower in processing visual information. Szucs et al. (2013) argue that the visual-spatial WM component is responsible for transformations of visual representations and operations that are relevant in mathematical problem solving. If more time is needed for these processes, it could be concluded that encoding information that is required for accurate transformation is hindered, and visual information cannot be properly processed in the visual-spatial workspace (Toll et al. 2016, p. 430f.). Since WM is considered a system of limited capacity (Baddeley 1983; Gathercole et al. 2016), our results match the points raised by Cornoldi and Giofrè (2014). They argue that response times could be interrelated with error rates in a way that longer response times might lead to losses of task-critical information that is held in temporary storage before being adequately manipulated or connected to long-term memory contents. Considering the role of fact retrieval from long-term memory for mathematical problem solving (e.g., Geary 1993) and the higher WM load in the case of using counting strategies instead of retrieving number facts from memory, we conclude that prolonged processing in WM might also affect mathematical performance.

A similar argument could be applied to account for longer response times in the verbal span task that we applied to assess verbal WM. Children with MLD needed significantly more time when they were asked to recall several numbers in a row. More numbers in mind might get lost when children need longer to process several numbers simultaneously in WM.

6.3 Differences in Mathematical Subtasks

With regard to the separate regressions on the different mathematical subtasks, we find accuracy and response time measures for verbal and visual-spatial WM tasks to be significantly related to all mathematical subtasks. This highlights WM to be an important domain-general cognitive resource for mental mathematical information. The fact that coefficients do not only become significant in tasks with an obvious relation to either verbal or visual-spatial type of information highlights that underlying cognitive processes of these relatively basic mathematical tasks are already multidimensional. Whenever we observe larger coefficients, they can be explained by the type of information which needs to be processed in the specific task. Verbal WM is particularly associated with the auditory arithmetic task. This might be due to the auditorily presented information in both, the verbal WM task and the auditory arithmetic task. This also applies to the central executive task, although the auditorily given information to be held active in the phonological loop are words instead of numbers. The need for visual-spatial processing in our written arithmetic task might not be obvious for this task with only written numbers as inputs and as results, but the high coefficient reflects visual-spatial processing to be relevant anyhow.

From research in mathematics education, it is well known that children with MLD typically have more difficulties switching between different forms of mathematical representations and that this skill is crucial in arithmetic problem solving (Gagatsis and Shiakalli 2004). In a similar manner, McCloskey et al. (1985) reasoned that number processing from verbally presented stimuli requires processing number words. Accordingly, successful translation depends on the child's ability to identify digits or numbers and to assemble the transmitted input into processable information. Moreover, the translation process involves the internalization of one external representation to produce another external representation (Pape and Tchoshanov 2001). If children fail to create an appropriate mental representation of the external one, this may result in an incorrect or slowed processing of information in WM already.

7 Conclusion and Outlook

Taken together, our results indicate that children with difficulties in learning mathematics do not only make more mistakes in WM tasks, but that they also need more time. This new result provides insights into the cognitive patterns underlying MLD.

7.1 Limitations

Nevertheless, we have to mention several limitations. It must be emphasized that WM and MLD are highly complex constructs that have to be simplified for the purpose of our research. The relationship between WM and MLD differs depending on age and on the specific WM and mathematical tasks (e.g., Menon 2016). For our study, this implies specific limitations: first, our phonological loop task was a forward digit span task. This is in line with the handling in the meta-analysis by Haberstroh and colleagues (DGKJP 2018a, p. 67), but the task is also used as a measure for short term memory in other literature (e.g., Bull and Johnston 1997) and backward span tasks are used more often as measures for the phonological loop. A stronger limitation results from the use of number as information to be remembered in this task. This leads to a confounding of content-specific processing difficulties of children with MLD. Hence, it is likely that the remarkably high effect size might thus be overestimated relative to e.g., a letter span task. Second, our three mathematical subtasks were methodologically restricted to tasks which could be solved on a touchscreen after a relatively short and highly standardized auditory instruction without any text and which could be rated as correct versus false. This remains an obvious limitation. Our tests only cover a small part of the multidimensional concept of mathematics. At the same time, peer studies cover slightly different parts in similar age groups (e.g., Klesczewski et al. 2018). Hence, comparisons or generalizations have to be taken with caution until more evidence is available. For example, we did not include mathematical word-problem solving tasks as in other studies in mathematics education (Fuchs et al. 2020; Verschaffel et al. 2020). Moreover, we did not include a task focusing on retrieval of number facts although these are also observed as a crucial difficulty in children with MLD

(e.g., Häsel-Weide 2016). This additional measure would have enabled us to better disentangle the effect of fact retrieval (with a low WM load) versus our tasks with a higher WM load as risk factors of MLD.

7.2 Outlook on Classroom Practice and on Further Research in Mathematics Education

Taking into consideration the previously mentioned evidence that WM is malleable by training and that this WM training can transfer to arithmetic and geometric performance in the long run (Berger et al. 2020; Judd and Klingberg 2021), our study highlights the importance of considering WM skills in the diagnosis and treatment of MLD. Following on from this, we discuss what can be derived for school practice and for future research in mathematics education.

First, including the diagnosis of WM skills and time dimensions in early digitally supported screenings might provide an opportunity to adapt the treatment more closely to the needs of the individual child. For example, if two children with comparably low mathematical skills differ in their WM capacity, the child with lower WM capacity might benefit more from a math-related digital WM training parallel to the MLD treatment than the other child with typical WM skills. The more detailed knowledge we have on the interrelations between WM and MLD, the more we can customize digitally supported treatment and prevention programs to the needs of individual children in inclusive classrooms.

Second, to ensure that WM does not become a limiting factor for children with MLD in mathematics learning, it is important to keep this cognitive resource in mind when planning lessons. On the one hand, our results emphasize the need for children with MLD to use visual mathematical aids (see e.g., Gaidoschik et al. 2021, p. 11) or (digitally generated) visual representations to reduce WM load and free WM capacities for mathematical learning (see Ladel 2020). Further practical implications to supporting children with MLD emerge when we consider our results in the context of recent evidence on the malleability of WM mentioned above: Adaptive digital WM exercises (Berger et al. 2020) or mathematical exercises on mental operations with visual representations of numbers (Wang et al. 2019) could be offered to children after diagnosing low mathematical and low WM skills to adaptively increase their WM capacity for future learning (see Winkel and Ladel 2022). The evidence by Wang and colleagues reminds us that our correlational results cannot simply be interpreted as causal evidence in one direction, but that WM skills and mathematical skills both can be mutually influential. Likewise, this raises new research questions about how we can design MLD treatments or everyday mathematics exercises to simultaneously promote WM and mathematical skills in children with MLD.

How exactly visual-spatial and verbal WM skills translate into mathematical skills or where exactly individual difficulties are anchored needs to be examined in depth in further interdisciplinary studies based on knowledge and methods from mathematics education, educational psychology and neurosciences. We see a high potential in the possibilities of digital tools to meet heterogeneous learning needs related to WM and mathematics (see also Winkel and Ladel 2022). Future research

is needed to develop and evaluate evidence-based learning environments to make this potential effective for children with MLD in inclusive classrooms.

References

Allen, K., Giofrè, D., Higgins, S., & Adams, J. (2020). Using working memory performance to predict mathematics performance 2 years on. *Psychological Research*. https://doi.org/10.1007/s00426-020-01382-5.

Baddeley, A. (1983). Working memory. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *302*(1110), 311–324. https://doi.org/10.1098/rstb.1983.0057.

Baddeley, A., & Hitch, G. (1974). Working memory. In G. Bower (Ed.), *Psychology of learning and motivation: advances in research and theory* (Vol. 8, pp. 47–89). Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1.

Benz, C., Grüßing, M., Lorenz, J.-H., Reiss, K., Selter, C., & Wollring, B. (2017). Zieldimensionen mathematischer Bildung im Elementar- und Primarbereich. In Stiftung „Haus der kleinen Forscher" (Ed.), *Frühe mathematische Bildung – Ziele und Gelingensbedingungen für den Elementar- und Primarbereich* (pp. 32–177). Barbara Budrich. https://doi.org/10.3224/84742051.

Berger, E., Fehr, E., Hermes, H., Schunk, D., & Winkel, K. (2020). The impact of working memory training on children's cognitive and Noncognitive skills. NHH Dept. of Economics discussion paper no. 09/2020. https://ssrn.com/abstract=3622985. Accessed 22.06.2023.

Bullheller, S., & Häcker, H. (2002). *Coloured Progressive Matrices (CPM)*. Deutsche Bearbeitung und Normierung nach JC Raven.

Bull, R., & Johnston, R. (1997). Children's arithmetical difficulties: Contributions from processing speed, item identification, and short-term memory. *Journal of Experimental Child Psychology*, *65*, 1–24. https://doi.org/10.1006/jecp.1996.2358.

Butterworth, B., & Laurillard, D. (2010). Low numeracy and dyscalculia: identification and intervention. *ZDM Mathematics Education*, *42*(6), 527–539. https://doi.org/10.1007/s11858-010-0267-4.

Cornoldi, C., & Giofrè, D. (2014). The crucial role of working memory in intellectual functioning. *European Psychologist*, *19*(4), 260–268. https://doi.org/10.1027/1016-9040/a000183.

De Weerdt, F., Desoete, A., & Roeyers, H. (2013). Working memory in children with reading disabilities and/or mathematical disabilities. *Journal of Learning Disabilities*, *46*(5), 461–472. https://doi.org/10.1177/0022219412455238.

DGKJP (2018a). S-3 Leitlinie: Diagnostik und Behandlung der Rechenstörung (Langfassung). AWMF-Leitlinien-Register-Nr. 028/046. https://register.awmf.org/assets/guidelines/028-046l_S3_Rechenst%C3%B6rung-2018-03_1.pdf. Accessed: 22.06.2023.

DGKJP (2018b). S-3 Leitlinie: Diagnostik und Behandlung der Rechenstörung (Leitlinienreport). AWMF-Leitlinien-Register-Nr. 028/046. https://register.awmf.org/de/leitlinien/detail/028-046. Accessed: 22.06.2023.

Dornheim, D. (2008). *Prädiktion von Rechenleistung und Rechenschwäche: der Beitrag von Zahlen-Vorwissen und allgemein-kognitiven Fähigkeiten*. Berlin: Logos.

Fischbach, A., Schuchardt, K., Brandenburg, J., Klesczewski, J., Balke-Melcher, C., Schmidt, C., Büttner, G., Grube, D., Mähler, C., & Hasselhorn, M. (2013). Prävalenz von Lernschwächen und Lernstörungen: Zur Bedeutung der Diagnosekriterien. *Lernen und Lernstörungen*. https://doi.org/10.1024/2235-0977/a000035.

Friso-Van den Bos, I., Van der Ven, S., Kroesbergen, E., & Van Luit, J. (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.

Fry, A., & Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology*, *54*, 1–34. https://doi.org/10.1016/S0301-0511(00)00051-X.

Fuchs, L., Fuchs, D., Seethaler, P., & Barnes, M. (2020). Addressing the role of working memory in mathematical word-problem solving when designing intervention for struggling learners. *ZDM Mathematics Education*, *52*(1), 87–96. https://doi.org/10.1007/s11858-019-01070-8.

Gagatsis, A., & Shiakalli, M. (2004). Ability to translate from one representation of the concept of function to another and mathematical problem solving. *Educational Psychology*, *24*(5), 645–657. https://doi.org/10.1080/0144341042000262953.

Gaidoschik, M., Moser Opitz, E., Nührenbörger, M., & Rathgeb-Schnierer, E. (2021). Besondere Schwierigkeiten beim Mathematiklernen. Special Issue der Mitteilungen der Gesellschaft für Didaktik der Mathematik, 47(111S). https://ojs.didaktik-der-mathematik.de/index.php?journal=mgdm&page=article&op=view&path%5B%5D=1042. Accessed: 22.06.2023.

Gathercole, S., Alloway, T., Willis, C., & Adams, A. (2006). Working memory in children with reading disabilities. *Journal of experimental child psychology*, *93*(3), 265–281. https://doi.org/10.1016/j.jecp.2005.08.003.

Gathercole, S., Woolgar, F., Kievit, R., Astle, D., Manly, T., Holmes, J., & Team, C. A. L. M. (2016). How common are WM deficits in children with difficulties in reading and mathematics? *Journal of Applied Research in Memory and Cognition*, *5*, 384–394. https://doi.org/10.1016/j.jarmac.2016.07.013.

Geary, D. (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin*, *114*(2), 345–362. https://doi.org/10.1037/0033-2909.114.2.345.

Gordon, R., Smith-Spark, J., Newton, E., & Henry, L. (2020). Working memory and high-level cognition in children: An analysis of timing and accuracy in complex span tasks. *Journal of Experimental Child Psychology*, *191*, 104736. https://doi.org/10.1016/j.jecp.2019.104736.

Graß, K., & Krammer, G. (2018). Direkte und indirekte Einflüsse der Raumvorstellung auf die Rechenleistungen am Ende der Grundschulzeit. *Journal für Mathematik-Didaktik*, *39*(1), 43–67.

Haberstroh, S., & Schulte-Körne, G. (2019). The diagnosis and treatment of dyscalculia. *Deutsches Ärzteblatt International*, *116*(7), 107–114M. https://doi.org/10.3238/arztebl.2019.0107.

Häsel-Weide, U. (2016). *Vom Zählen zum Rechnen: struktur-fokussierende Deutungen in kooperativen Lernumgebungen*. Springer Spektrum. https://doi.org/10.1007/978-3-658-10694-2.

Jordan, N.C., Glutting, J., & Ramineni, C. (2010). The importance of number sense to mathematics achievement in first and third grades. *Learning and individual differences*, *20*(2), 82–88. https://doi.org/10.1016/j.lindif.2009.07.004.

Judd, N., & Klingberg, T. (2021). Training spatial cognition enhances mathematical learning in a randomized study of 17,000 children. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-021-01118-4.

Kaufmann, L., & von Aster, M. (2012). The diagnosis and management of dyscalculia. *Deutsches Ärzteblatt international*, *109*(45), 767–778. https://doi.org/10.3238/arztebl.2012.0767.

Klesczewski, J., Brandenburg, J., Fischbach, A., Schuchardt, K., Grube, D., Hasselhorn, M., & Büttner, G. (2018). Development of working memory from grade 3 to 5: Differences between children with and without mathematical learning difficulties. *International Journal of Disability, Development, and Education*, *65*, 509–525. https://doi.org/10.1080/1034912X.2017.1419555.

Ladel, S. (2020). Bildungstechnologie im Mathematikunterricht (Klassen 1–6). In *Handbuch Bildungstechnologie* (pp. 645–666). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-662-54373-3.

Landerl, K., Bevan, A., & Butterworth, B. (2004). Developmental dyscalculia and basic numerical capacities: A study of 8–9-year-old students. *Cognition*, *93*, 99–125. https://doi.org/10.1016/j.cognition.2003.11.004.

Maehler, C. (2021). Diagnostik von Lernstörungen: Zeit zum Umdenken. *Zeitschrift für pädagogische Psychologie*, *35*(4), 217–227. https://doi.org/10.1024/1010-0652/a000291.

Maehler, C., & Schuchardt, K. (2016). Working memory in children with specific learning disorders and/or attention deficits. *Learning and Individual Differences*, *49*, 341–347. https://doi.org/10.1016/j.lindif.2016.05.007.

McCloskey, M., Caramazza, A., & Basili, A. (1985). Cognitive mechanisms in number processing and calculation: Evidence from dyscalculia. *Brain and Cognition*, *4*, 171–196. https://doi.org/10.1016/0278-2626(85)90069-7.

Menon, V. (2016). Working memory in children's math learning and its disruption in dyscalculia. *Current Opinion in Behavioral Sciences*, *10*, 125–132. https://doi.org/10.1016/j.cobeha.2016.05.014.

Murphy, M., Mazzocco, M., Hanich, L., & Early, M. (2007). Cognitive characteristics of children with mathematics learning disability (MLD) vary as a function of the cutoff criterion used to define MLD. *Journal of Learning Disabilities*, *40*(5), 458–478. https://doi.org/10.1177/00222194070400050901.

Pape, S., & Tchoshanov, M. (2001). The role of representation(s) in developing mathematical understanding. *Theory Into Practice*, *40*(2), 118–127. https://doi.org/10.1207/s15430421tip4002_6.

Passolunghi, M., & Costa, H. (2019). Working memory and mathematical learning. In *International handbook of mathematical learning difficulties* (pp. 407–421). Cham: Springer. https://doi.org/10.1007/978-3-319-97148-3_25.

Peijnenborgh, J., Hurks, P., Aldenkamp, A., Vles, J., & Hendriksen, J. (2016). Efficacy of working memory training in children and adolescents with learning disabilities: A review study and meta-analysis. *Neuropsychological rehabilitation*, *26*(5), 645–672. https://doi.org/10.1080/09602011.2015.1026356.

Peng, P., & Fuchs, D. (2016). A meta-analysis of working memory deficits in children with learning difficulties: Is there a difference between verbal domain and numerical domain? *Journal of Learning Disabilities*, *49*(1), 3–20. https://doi.org/10.1177/0022219414521667.

Ranger, J., & Kuhn, J.-T. (2012). A flexible latent trait model for response times in tests. *Psychometrika*, *77*(1), 31–47. https://doi.org/10.1007/s11336-011-9231-7.

Rousselle, L., & Noël, M. (2007). Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs non-symbolic number magnitude processing. *Cognition*, *102*, 361–395. https://doi.org/10.1016/j.cognition.2006.01.005.

Ruwisch, S., & Lorenz, J.-H. (2018). Entstehung der Leitlinie zur „Diagnostik und Behandlung der Rechenstörung" und Beteiligung der GDM. *Mitteilungen der Gesellschaft für Didaktik der Mathematik*, *44*(105), 41–44.

Scherer, P., Beswick, K., DeBlois, L., Healy, L., & Moser Opitz, E. (2017). Assistance of students with mathematical learning difficulties—how can research support practice? A summary. In *Proceedings of the 13th International Congress on Mathematical Education* (pp. 249–259). Cham: Springer. https://doi.org/10.1007/978-3-319-62597-3_16.

Schindler, M., Schovenberg, V., & Schabmann, A. (2020). Enumeration processes of children with mathematical difficulties: an explorative eye-tracking study on subitizing, groupitizing, counting, and pattern recognition. *Learning Disabilities: A Contemporary Journal*, *18*(2), 193–211.

Schipper, W. (2002). Thesen und Empfehlungen zum schulischen und außerschulischen Umgang mit Rechenstörungen. *Journal für Mathematik-Didaktik*, *23*, 243–261. https://doi.org/10.1007/BF03338958.

Schneider, W., Küspert, P., & Krajewski, K. (2021). *Die Entwicklung mathematischer Kompetenzen* (3rd edn.). Stuttgart: UTB.

Schuchardt, K., Maehler, C., & Hasselhorn, M. (2008). Working memory deficits in children with specific learning disorders. *Journal of Learning Disabilities*, *41*(6), 514–523. https://doi.org/10.1177/0022219408317856.

Szucs, D., Devine, A., Soltesz, F., Nobes, A., & Gabriel, F. (2013). Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. *Cortex*, *49*, 2674–2688. https://doi.org/10.1016/j.cortex.2013.06.007.

Toll, S. W., Kroesbergen, E., & Van Luit, J. (2016). Visual working memory and number sense: Testing the double deficit hypothesis in mathematics. *British Journal of Educational Psychology*, *86*, 429–445. https://doi.org/10.1111/bjep.12116.

Verschaffel, L., Schukajlow, S., Star, J., & Van Dooren, W. (2020). Word problems in mathematics education: A survey. *ZDM Mathematics Education*, *52*(1), 1–16. https://doi.org/10.1007/s11858-020-01130-4.

Viesel-Nordmeyer, N., Ritterfeld, U., & Bos, W. (2020). Welche Entwicklungszusammenhänge zwischen Sprache, Mathematik und Arbeitsgedächtnis modulieren den Einfluss sprachlicher Kompetenzen auf mathematisches Lernen im (Vor-) Schulalter? *Journal für Mathematik-Didaktik*. https://doi.org/10.1007/s13138-020-00165-0.

Von Aster, M., & Lorenz, J.-H. (2013). *Rechenstörungen bei Kindern: Neurowissenschaft, Psychologie, Pädagogik*. Vandenhoeck & Ruprecht.

Wang, C., Xu, T., Geng, F., Hu, Y., Wang, Y., Liu, H., & Chen, F. (2019). Training on abacus-based mental calculation enhances visuospatial working memory in children. *Journal of Neuroscience*, *39*(33), 6439–6448. https://doi.org/10.1523/JNEUROSCI.3195-18.2019.

WHO (World Health Organization) (2022). International statistical classification of diseases and related health problems. Eleventh revision. (ICD-11), version 02/2022. https://icd.who.int/browse11/l-m/en. Accessed: 22.06.2023.

Winkel, K., & Ladel, S. (2022). Heterogenen Lernvoraussetzungen digital Gerecht werden – Digitale Werkzeuge zur Unterstützung des Arbeitsgedächtnisses im Mathematikunterricht der Grundschule. In B. Brandt, L. Bröll & H. Dausend (Eds.), *Digitales Lernen in der Grundschule III. Fachliche Lernprozesse im Elementar- und Primarbereich anregen* (pp. 381–393). Münster: Waxmann. https://doi.org/10.31244/9783830995913.