Learning Disabilities: A Contemporary Journal 18(2), 193-211, 2020

Enumeration Processes of Children With Mathematical Difficulties: An Explorative Eye-Tracking Study on Subitizing, Groupitizing, Counting, and Pattern Recognition

Maike Schindler*
Vanessa Schovenberg
Alfred Schabmann
University of Cologne

This article investigates how students with mathematical difficulties (MD) differ from typically developing (TD) students in enumeration processes of small sets of objects (of 1 up to 9 dots). We present a study with 20 fifth-grade students of which ten were found to have MD in initial diagnostics. The students were supposed to exactly enumerate sets of dots, i.e., to say how many dots they saw. This took place in three conditions: (a) in random arrangements in the subitizing range (1–4 dots), (b) in random arrangements in the counting range (5–9 dots), and (c) in canonical (dice-like) arrangements (1–9 dots). We used eye tracking (ET) to analyze student enumeration processes derived from ET video data. Whereas we did not find significant group differences in students' error rates, we found differences in response times with longer response times for students in MD in the canonical arrangement condition. Further, we found significant group differences in students' enumeration processes in all three conditions (subitizing range, counting range, canonical): Students with MD tended to count all dots more often whereas TD students used more advantageous enumeration processes such as simultaneous enumeration or enumeration of groups of dots more often. Our results support the assumption of qualitatively different enumeration processes between students with and without MD.

Keywords: **Mathematical Difficulties, Enumeration, Subitizing, Counting, Eye tracking**

Introduction

The ability to enumerate small sets of objects, i.e., to grasp a set of 1 up to 9 dots and to say how many they are, is significant for children to learn basic arithmetic, and hence is critical for students' mathematical development (Starkey, & Cooper, 1995). Exact enumeration of small sets can involve different processes: Small sets of objects can be enumerated through subitizing, i.e., the recognition of sets of items up to 3 or 4 at a glance, which is an automatized perception process (Clements, 1999), but also through counting (Gelman & Gallistel, 1986), or through perceiving sets in subgroups (so-called "groupitizing", Starkey & McCandliss, 2014). Further, enumeration can involve pattern recognition, where certain familiar patterns (e.g., dice patterns) are recalled and recognized at a glance (Ashkenazi, Mark-Zigdon, & Henik, 2013).

*Please send correspondence to: Maike Schindler, PhD, Junior Professor, Department of Special Education and Rehabilitation, University of Cologne, Classen-Kappelmann-Str. 24, 50931 Köln, Germany, Phone: +49 (0) 221 470-76726, Email: maike.schindler@uni-koeln.de.

However, research indicates that not all people can enumerate small sets of objects in the same way: For example, the subitizing mechanism appears not to function properly for students with MD (Landerl, 2013; Schleifer & Landerl, 2011; Van der Sluis et al., 2014, see Section on Enumeration in more detail). Given the significance of subitizing for the development of the number concept and for arithmetic learning (Starkey & Cooper, 1995), difficulties in subitizing constitute a risk factor for students' mathematics learning. Mock et al. (2016)—referring to previous research (Moeller et al., 2009)—summarize that MD[1] "may indeed emerge (among others) from a deficit in subitizing and thus, the automatic and parallel enumeration of small set sizes" (p. 337).

Research into enumeration processes of students with MD so far predominantly investigates students' response times. However, whether students with MD are slower in subitizing or whether enumeration processes are qualitatively different (with other enumeration processes for students with MD than for TD students) can hardly be explained through response times data: Van der Sluis, de Jong, and van der Leij (2014) conclude from their study that "[w]hether the arithmetic-disabled children in our study were slower in subitizing, or whether they (occasionally) reverted to a safer counting strategy, cannot be determined with the current data" (p. 260), since response times were studied.

Addressing this research gap, a case study (Moeller et al., 2009) has started to use eye tracking (ET), the recording of students' eye movements, to further investigate group differences between students with and without MD on enumeration of small sets—with promising insights into enumeration processes of two children with MD. These insights were the starting point for our study: The aim of this explorative study is to investigate if students with and without MD differ in enumeration processes of small sets of objects (1–9) on group level. For pursuing this aim, we use ET to inquire into student enumeration processes and see if students with MD use different enumeration processes for small sets of items than their TD peers.

Mathematical Difficulties (MD)

Depending on different national contexts, disciplines, and studies, researchers speak of mathematical difficulties, mathematical learning disabilities or developmental dyscalculia (Moser Opitz et al., 2016; Scherer et al., 2016). The ICD-10 describes mathematical difficulties as a developmental learning disorder with impairment in arithmetical skills that cannot be explained by a general mental retardation or inadequate schooling (WHO, 2019). This supports an IQ-discrepancy-model, where a disorder in mathematics is diagnosed based on a discrepancy between the IQ and the mathematics performance of the person. However, recent research has indicated that both groups of people (those where mathematics and intelligence scores differ substantially from each other and those where they don't) show similar patterns in mathematical problem solving, in particular they "do not show qualitatively different cognitive patterns in counting, subitizing, or magnitude comparison" (Kuhn et al., 2013, p. 244). This suggests that a group separation based on the discrepancy criterion is unrewarding—at least when basic numerical abilities are assessed (ibid, p.

1 Even though some authors, such as Mock et al. (2016), use the term *developmental dyscalculia*, our article uses the term *mathematical difficulties (MD)*. See Section on Mathematical Difficulties.

242). In our research, we use the term *mathematical difficulties (MD)* and—following Moser Opitz et al. (2016) and Scherer et al. (2016)—address students with difficulties in, for example counting by groups and counting principles, the base-10 number system and understanding the place value, understanding of operation of multiplication, division and supplementing (see also Simon & Grünke, 2010).

Enumeration

Enumeration of small sets: Subitizing, groupitizing, counting, and pattern recognition

Our research focuses on students' ability to enumerate the exact number of 1 up to 9 items, which predominantly involves subitizing, counting, and groupitizing (Ashkenazi et al., 2013, Starkey & McCandliss, 2014). Besides this, enumeration can involve *estimation,* "an imprecise process used to assess large numbers of items for a short presentation time" (Ashkenazi et al., 2013, p. 36). There is whole body of research on students' estimation of larger sets of dots, connected to the concept of Approximate Number System (ANS). In this line of research, participants are asked to verbally estimate the number of sets of dots (or to say which of two sets of dots is greater) while the dots are presented only shortly to prevent exact enumeration and counting (Bugden & Ansari, 2016; Mazzocco, Feigenson, & Halberda, 2011). Yet, the imprecise estimation of larger sets was not in the focus of our study and research on the ANS is beyond the focus of our article. Therefore, we do not go into detail about the ANS. Instead, our article falls within the body of research focusing on enumeration processes connected to the exact enumeration of smaller sets, as outlined in the following.

Subitizing—a word derived from the latin word subitus (suddenly)—is understood as the ability to fast and exactly name the numbers of small quantities without counting (Fischer, Gebhardt, & Hartnegg, 2008). Researchers currently agree that humans are able to perceive quantities up to 4 (Clarke, Cheeseman, & Clarke, 2006; Mandler & Shebo, 1982; Schleifer & Landerl, 2011). Subitizing is assumed to be a perceptual, automatized, and often subconscious process (Clements, 1999). Even young children at the age of two are able to subitize perceptually (e.g., Starkey & Cooper, 1980, 1995). Subitizing constitutes the basis for the development of the number concept and is an important requirement for arithmetic learning (ibid).

Enumeration can also involve patterning abilities, when students perceive sets in subsets (Clements, 1999). Such *groupitizing* (Starkey & McCandliss, 2014; or "conceptual subitizing", Clements, 1999) involves abilities such as composing and decomposing as well as understanding the concepts of number and part-whole schema (Starkey & McCandliss, 2014), i.e., the "understanding that quantities can be decomposed into pieces and reassembled again" (Krajewski & Schneider, 2009, p. 513), which in most cases children already at a young age use flexibly in arithmetic. In particular, groupitizing requires the insight that numbers are compounded by other numbers, for example, that 7 is a compound of 4 and 3, etc. The representation of items in groups appears to help school students to enumerate sets of items (Starkey & McCandliss, 2014).

Quantities greater than 4 can (in most cases) not be perceived at a glance—they need to be counted (or perceived in groups). *Counting*, as compared to subitizing, is a "slower and more error-prone process of one-to-one mapping between a set of objects and number words" (Schleifer & Landerl, 2011, p. 280; see also Gelman & Gallistel, 1986). Scholars speak of the *counting range* with respect to sets with 5 or more items, and refer to the *subitizing range* for sets up to 4 items (e.g., Ashkenazi et al., 2013; Schleifer & Landerl, 2011).

Enumeration can also involve the recognition of patterns. If sets are, for instance, arranged in dice-like or domino-like patterns, so-called *canonical arrangements*, quantities can be recalled and recognized when the pattern and the according number are present to the child (e.g., Starkey & McCandliss, 2014). In this case, the child then does not actually pay attention to every single item (dot) but associates the very pattern to a number word (Von Glasersfeld, 1982).

Related work: Enumeration performance of students with MD

Scholars have studied exact enumeration processes among students with MD mostly without the use of ET. For example, Fischer et al. (2008), investigating 375 students aged 7–17, found that the capacity to enumerate briefly presented items is lower for students with MD than for TD students, even in the subitizing range. Landerl (2013), in a similar vein, found in a study on 83 students in grades 2 to 4 that students with MD had longer response times, and particularly larger slopes in the subitizing range. Similarly, Schleifer and Landerl (2011), investigating 52 students with MD attending grades 2 to 4 and a control group of 52 TD children, found that students with MD showed steeper response times slopes in the subitizing range, while the response times slopes in the counting range were comparable between the groups—indicating that the subitizing mechanism does not function properly for students with MD (Schleifer, & Landerl, 2011). Gray and Reeve (2014), investigating dot enumeration of 78 preschool children, found that weak subitizing profiles were related to poor addition abilities, indicating poor subitizing abilities for students with MD. Further, van der Sluis et al. (2014), in a study with 74 children in grades 4 to 5 found that students with low mathematical abilities were slower in naming quantities in the subitizing range. These findings all indicate particular subitizing difficulties for students with MD (see Landerl, 2013). Differences in enumeration between students with and without MD also apply canonical representations: Ashkenazi et al. (2013), investigating eleven children with MD and a control group of eleven TD children, found that students with MD were less accurate than the control group and that accuracy decreased as the number of dots increased.

One study has furthermore used ET to investigate group differences between students with and without MD: Moeller et al. (2009) investigated two ten-year-old boys with MD and eight TD children and compared their abilities of enumerating small sets of 1 to 8 items. They observed that children with MD differed from the control group: While the two boys with MD counted, the TD students enumerated certain quantities in parallel or rather simultaneously. In the subitizing range, response times of students with MD were also higher. Furthermore, one boy with MD showed more fixations (i.e., relatively still eye movements), whereas the other boy had a longer fixation duration on average in the counting range as compared to the TD students.

This study

The aim of this explorative study is to investigate if students with and without MD differ in enumeration processes of small sets of objects (1–9) on group level. Among others, we use ET, the recording of eye movements, since it has the potential to offer insights into student enumeration processes, but avoids a verbalization step, allowing immediate access to enumeration processes. Previous research indicates that ET data represent a valuable and promising supplement to current diagnostic measures for investigating MD (Mock et al., 2016; Schindler et al., 2019; Van Viersen et al., 2013). Using ET, we seek to identify student enumeration processes: Such processes are, for example, simultaneous perception, using groups or structures to enumerate, or counting serially.

We ask the research question: *Do students with and without MD differ in dot enumeration, in particular in (1) error rates, (2) response times, (3) numbers of fixation, and (4) enumeration processes as inferred from eye movements?* We do so by investigating *random arrangements* in (a) the subitizing range (1–4 items) and (b) the counting range (5–9 items), and (c) *canonical arrangements* (of 1–9 items).

Method

Participants

To investigate group differences in enumeration processes of students with and without MD, we conducted an explorative study and use data from 20 fifth-grade students from a German inclusive comprehensive school with medium school economic status[2] (see Table 1). The study took place in the first weeks of the school year, after the students had finished primary school.

For recruiting the participants of our study, a pre-selection was made by the teachers: The teachers named students who they thought might have MD and others who they thought were not having MD. Based on the teachers' guidance, 23 students were asked to participate, whereof 20 were willing to participate and—in the following—took part in standardized testing regarding their arithmetic performance, to identify if they had MD or were TD (see below).

All students were taught the same mathematical curriculum and were schooled in whole-class mathematics teaching. At the time of this study, students who participated in this study did not receive any supplemental mathematics instruction.

Since the students came from different primary schools and had attended their new school only for some weeks, information about prior mathematics teaching methods was not available. Yet, all students came from the same state, so the same mathematical curriculum applied. Four students had special educational needs—in social and emotional development (2), physical development (1), and learning (1). All of them had been schooled in inclusive primary schools before in regular mathematics classes. Even the student with special needs in learning had participated

2 In this part of Germany, the schools are assigned "location types" based on the students' achievements in mathematics, German, and English, the ratio of students with immigration background, and the percentage of unemployment/welfare-collection among legal guardians. The "location types" are ranked 1 (favorable preconditions) to 5 (unfavorable preconditions). The participating school was ranked 3 (medium).

in regular mathematics classes. Though he received special support in primary school, information on the contents of this support was not available.

Arithmetic performance. A standardized arithmetic paper-pencil speed test was administered to all 20 students: HRT (Haffner et al., 2005). HRT is frequently used to diagnose mathematical difficulties in German speaking countries—also in studies investigating enumeration of small sets of objects (e.g., in Moeller et al., 2009; Schleifer & Landerl, 2011). HRT diagnoses MD at a percentage rank (PR) ≤10. It has two parts: For diagnosing MD, either the full test or only the first part, focusing on number and arithmetic, can be administered (Haffner et al., 2005). For the purpose of this study, only the first part (the arithmetic part) was administered (similar to Landerl, 2013; Schleifer & Landerl, 2011), which has a high reliability (r_{tt}=.93) and high correlation with other standardized mathematics tests (*r*=.72, Haffner et al., 2005). The norm sample of *N*=3,354 students in Germany matched the sample in our study.

The test was administered by the first author of this article. It is a class test: All students worked on the tasks in a booklet at the same time. In the beginning, general information was given to the students, based on the standardized instructions of the test. Every subtest contains a list of tasks with slightly increasing difficulty. The time on each subtest was 2 mins: In that time, the students were asked to write down as many correct answers as possible. Four subtests address arithmetic operations (addition, subtraction, multiplication, and division; e.g., 5 + 3 = _ or 3 x 5 = _). The two further subtests contain completion tasks (e.g., _ - 2 = 6) and symbolic quantity comparison (e.g., 18 _ 7 [correct response: >]). All subtests were administered in one sitting.

Based on the results from HRT, eight students had MD (at PR≤10), nine students were TD (at PR>25), and three students were assigned a PR of 14 and 16 respectively, which is within the "at risk zone" (Haffner et al., 2005, p. 20). We also conducted *qualitative diagnostic interviews* (following Wartha & Schulz, 2012) with all students with PR≤25 (i.e., 11 students). The analysis revealed that one of the students in the at-risk zone did not show MD and thus was counted to the TD group (the other two showed MD). Thus, both groups finally had ten students.

Visuo-spatial working memory. Since deficits in visuo-spatial working memory can affect the subitizing ability (Peterson & Simon, 2000), we examined students' visuo-spatial working memory using two subtests of the computer-based AGTB 5–12 (Hasselhorn et al., 2012), a working memory test battery. AGTB was standardized based on a sample of *N*=1,658 students in Germany, which matched the sample of our study. The *Matrix* task addresses the visuo-static component of the visuo-spatial working memory ("visual cache"). Here, the participants see a grid of 4x4 white fields. In each item, a static pattern of 2–8 black fields (other fields: white) appears. The participants are to remember this pattern and recall it. The *Corsi Block* task addresses the visuo-dynamic component of the visuo-spatial working memory ("inner scribe"[3]). Here, nine white fields are scattered on a grey background. In each

3 "The visual cache stores information primarily about visual form and color and is closely linked to the visual perceptual system. The inner scribe retains information about movement sequences and is closely linked to the planning and execution of movement." (Logie & Pearson, 1997, p. 243)

item, a smiley appears on different fields, one after another, in a certain "path". The participants are supposed to recall this path (positions and order of the smileys) correctly.

Table 1. ***Participants of the study. Means (SDs) of descriptive measures and performances in achievement tests***

	MD group (n=10)	TD group (n=10)
Participant information		
Age	10;10 (0;6)	10;7 (0;5)
Gender (% girls)	60.0	40.0
Special needs (%)	40.0	00.0
Mother tongue (% German)	70.0	80.0
Performances (in -scores)		
Arithmetic (HRT)	-2.1 (0.7)	-0.1 (0.6)
Visuo-spatial working memory (AGTB): – Visual Cache ("Matrix")	0,0 (0.5)	-0,2 (0.2)
– Inner Scribe ("Corsi Block")	-0,3 (0.4)	-0,4 (0.6)

Apparatus

We recorded eye movements with wearable ET glasses: Tobii Pro Glasses 2. For the explorative purpose of our study, we decided not to use a screen-based eye-tracker (similar to, e.g., Schindler et al., 2019; Schindler & Lilienthal, 2018). Screen-based eye trackers, connected to the monitor, have certain advantages, especially since the analysis of ET measures (such as numbers of fixations) is less laborious. However, in our explorative study we wanted to investigate students' gazes in more detail. We wanted to observe, for instance, when students' gazes went away from the screen (and where exactly). A screen-based eye tracker would not have tracked these gazes. The use of glasses further allowed us to record the video of the scene view, ET and audio data in a time-synchronized manner.

Tobii Pro Glasses 2 have a high-resolution scene camera, ET sensors, and a microphone. The scene camera records an HD video stream (1920 x 1080 pixels) of the participants' view with a field of view of 82° (horizontal) x 52° (vertical). Infrared illuminators and ET sensors (cameras pointing towards the participant) allow to track the eye gaze with a sampling rate of 50 Hz by identifying the displacement of corneal reflections from the detected center of the pupil.

The recording unit of the Pro Glasses 2 records and saves the ET data. With a weight of 45 gr, the glasses are comfortable and enable a relative freedom of movement. The stimuli were displayed on a 24" HD monitor with a refresh rate of 60 Hz. The accompanying controller software of the Pro Glasses 2 enables the management of the data during the test procedure. It allows the calibration of the glasses and the administration of the recordings.

Procedure, tasks, and analysis

The individual experiments took place in a quiet room in the students' school. The students sat still on a firm chair, with a viewing distance of approx. 50 cm to the monitor, and a stable viewing angle. The experimenter (first author of this paper) controlled the notebook, which was connected to the screen (for displaying the tasks). For adjusting the eye tracker, a one-point calibration was conducted. After a first acquaintance phase with the monitor and eye tracker, the students were introduced to the enumeration tasks.

Computerized enumeration tasks were presented. In each item, the students were presented red dots (1–9 dots) in the center of the white screen (dot size on the 24" monitor: 2 cm, maximum span (vertical/horizontal): 15 cm). A total of 36 items were presented. The quantities (1–9) were each presented in four different arrangements, once in a canonical arrangement (Fig. 1), and in three random arrangements, arranged differently each time (see Table 2 for examples).

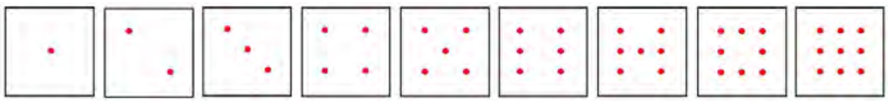


Figure 1. **Canonical arrangement of dots**

The students were instructed to say the number of dots that was presented correctly and as fast as possible. The trial started with three practice items to ensure that the participant understood the task. In between the items, a white screen with a fixation-star (similar to a fixation-cross) appeared to ensure a clear transition from one item to the next. The students were asked to fixate the star before the next item appeared. The canonical and randomly arranged items were intermixed randomly. The students did not receive information about whether their answers were correct or incorrect. *Response times* were recorded from appearance of the dot-item until the given answer. Verbal answers were recorded to account for *error rates. Numbers of fixations* were used as output of Tobii Pro Lab software. For analyzing *student enumeration processes*, we used gaze-overlaid videos provided by Tobii Pro Lab software: videos with eye gazes represented as a semi-transparent dot wandering around in the video. For 640 items (20 students x 32 items), we analyzed gaze-overlaid videos to infer student enumeration processes. We analyzed the data in three stages. *First stage:* We assigned enumeration processes to the clips of gaze-overlaid videos inductively, following Mayring's (2014) qualitative content analysis in an inductive manner. For instance, when nine dots were shown, and a student looked at every dot one after the other, we assigned the process "counting all." We did so for all three conditions: (a) the subitizing range in random arrangements, (b) the counting range in random arrangements, and (c) the canonical arrangements. *Second stage:* We refined the set of categories in a *category revision* step. This revision step resulted in three categories of processes for each condition, describing the processes that the students used to enumerate quantities (Table 2).

Third stage: Based on the set of categories, all gaze-overlaid video data of all students were coded. For every item, a category was allocated. Following Mayring (2014) and Döring and Bortz (2016), 25 % of the data were coded by two raters independently to investigate interrater reliability. Both raters were researchers in the field, had experience with gaze-overlaid video analysis, and used the list of categories with according descriptions of the categories. We calculated the interrater reliability using Cohen's kappa (Cohen, 1988). The interrater agreement was 0.93, which can be considered excellent or almost perfect (Landis & Koch, 1977).

Statistical analysis

For the statistical analysis, the statistics and analysis software IBM SPSS 26 was used.

Error rates: The total number of errors was relatively low. First, we conducted a Shapiro-Wilks Test for normality (Shapiro & Wilk, 1965). Based on the result of this test, we found that error rates were not normally distributed (Tab. 3). Thus, we used the Mann-Whitney U test as non-parametric test suitable for small sample sizes.

Response times/Numbers of fixations: For comparing the groups (students with MD vs. TD students) with respect to response times/numbers of fixations, we carried out repeated measurement ANOVAS with type of measure (subitizing, counting, canonical condition) as within factor and group (MD, TD) as between factor.

Student enumeration processes: To identify group differences in student enumeration processes, we carried out chi square tests based on the total of enumeration processes used by the students. Chi square tests were carried out separately for subitizing, counting, and the canonical condition.

Results

Error rates

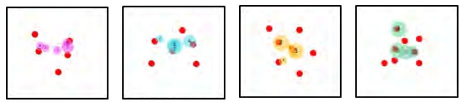
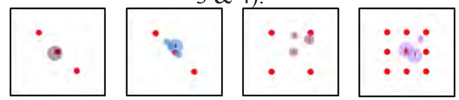
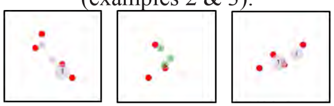
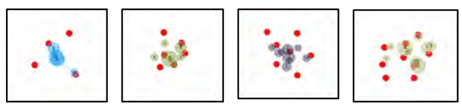
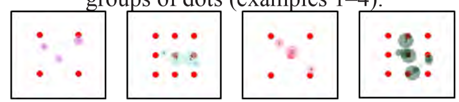
In line with previous studies (e.g., Mandler & Shebo, 1982; Piazza et al., 2002; Schleifer & Landerl, 2011), the total number of errors was generally low (Table 3). In 720 items, the students made 27 errors (3.75%) overall. While students with MD had an error rate of 4.72 % (17 errors/360 items), the error rate of TD students was 2.77 % (10 errors/360 items). Despite the slightly higher error rates for the group of students with MD, group differences were not significant (U = 36.00, p = .32, d = 0.49; 95%-CI [-0.464; 1.444]).

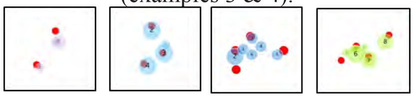
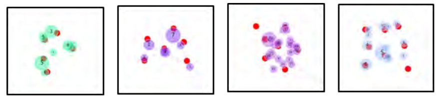
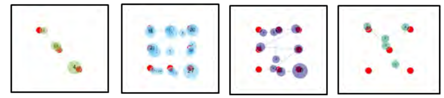
Table 3. *Errors. Comparison of groups. Mean (SD), skewness, kurtosis, Shapiro-Wilks Test for normality (W)*

MD group				TD group			
M (SD)	skewness	kurtosis	*W*	*M (SD)*	skewness	kurtosis	*W*
1.70 (1.77)	1.78	3.78	.81*	1.00 (1.33)	1.41	1.74	.79*

Note: * p < .05

Table 2. ***Categories of student enumeration processes—derived from gaze-overlaid videos***

Random Arrangement		Canonical Arrangement (2–9)
Subitizing Range (2–4)	Counting Range (5–9)	
1) Simultaneous enumeration Presented dots are enumerated simultaneously: Either (a) there is no eye movement towards a dot at all but the gaze remains in the position of the previous fixation cross (examples 1 & 2) or (b) the gaze moves to the center of the items (examples 3 & 4).  Example 1 Example 2 Example 3 Example 4	**1) Quasi-simultaneous enumeration** Presented dots are enumerated quasi-simultaneously: Gazes are on the groups of dots, whereas single dots are perceived peripheral (examples 1–4).  Example 1 Example 2 Example 3 Example 4	**1) Simultaneous enumeration** Presented dots are enumerated simultaneously: Either (a) there is no eye movement, only the gaze in the middle (example 1), (b) there are eye movements only in the center of the item (example 2), or (c) there are gazes to only one side, to one corner or to another dot (examples 3 & 4).  Example 1 Example 2 Example 3 Example 4
2) Enumeration of groups of dots Presented dots are enumerated in groups: The gazes do not go to every single dot presented, but (a) to groups of dots (example 1), or (b) to one group with one remaining dot (examples 2 & 3).  Example 1 Example 2 Example 3	**2) Partial enumeration of groups of dots** Presented dots are partially enumerated in groups: For parts of the presented dots, the gazes are on groups of dots, for other parts, the gazes go to every dot (examples 1–4).  Example 1 Example 2 Example 3 Example 4	**2) Enumeration through use of groups / structures** Presented dots are enumerated through the use of groups / structures: Gazes are on parts of the dots, indicating the use of symmetries or groups of dots (examples 1–4).  Example 1 Example 2 Example 3 Example 4

3) Counting all Presented dots are counted serially: The gazes (a) go to every dot (examples 1 & 2) or (b) in case of four dots, they go to three out of four (examples 3 & 4).  Example 1 Example 2 Example 3 Example 4	**3) Counting all** Presented dots are counted serially: The gazes (a) go to every dot (examples 1 & 2) or (b) go to all dots but one (examples 3 & 4).  Example 1 Example 2 Example 3 Example 4	**3) Counting all** Presented dots are counted serially: The gazes (a) go to every dot (examples 1 & 2) or (b) go to all dots but one (examples 3 & 4).  Example 1 Example 2 Example 3 Example 4

Response times

A 2 (group: students with MD vs. TD students) × 3 (measure: subitizing range, counting range, canonical arrangement) repeated measurement ANOVA revealed significant effects for measure (F (2, 36) = 160.59, p = .000, η_p^2 = .89 (95%-CI: [.82; .93])) and the group × measure interaction (F (2, 36) = 3.68, p = .035, η_p^2 = .17 [.00; .35]). No significant main effect for group was found. Inner subject contrasts revealed significant differences between the counting range condition and the other conditions for the measure effect (F (1, 18) = 207.54, p = .000, η^2 = .92 [.81; .95]). The contrast for the interaction was significant for the canonical condition vs. all others (F (1, 18) = 7.56, p = .014, η^2 = .29 [.01; .54]). Figure 2 (left) reveals that only in the canonical condition, group differences in response times occur (t (18) = 2.55, p = .020, d = 1.14 [0.13; 2.15]).

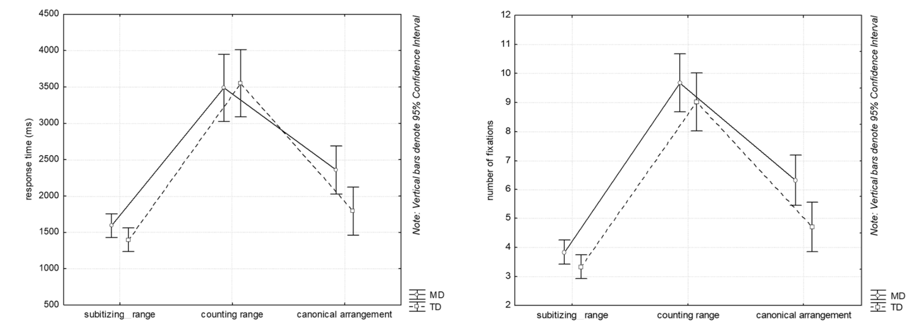


Figure 2. **Response times (left) and numbers of fixations (right) in three conditions**

Numbers of fixations

A 2 (group: students with MD vs. TD students) × 3 (measure: subitizing range, counting range, canonical arrangement) repeated measurement ANOVA revealed significant effects for measure (F (2, 36) = 245,45, p = .000, η_p^2 = .93 [.88; .95]), but no significant group effect and no significant interaction. Inner subject contrasts revealed significant differences between the counting range condition and the other conditions for the measure effect (F (1, 18) = 298.17, p = .000, $\eta 2$ = .94 [.86; .97]) (see Figure 2).

Student enumeration processes

Chi square tests with the accumulated processes revealed significant differences in the distribution of enumeration processes between the two groups for all the three conditions (subitizing range: χ^2(2) = 17.95, p = .000; counting range: χ^2(2) = 57.3, p = .000; canonical arrangement: χ^2(2) = 11.85, p = .000). Cell tests for the differences between groups revealed the following results: *Subitizing range:* Students with MD used the counting all process more often than TD students (χ^2(1) = 4.83, p = .028). *Counting range:* Students with MD used less often the quasi-simultaneous enumeration (χ^2(1) = 6.49, p = .010) and the partial enumeration of groups of dots (χ^2(1) = 5.53, p = .018), but more often the counting all process

($\chi^2(1) = 17.119$, $p = .000$) than TD students. *Canonical arrangement:* Here, singular cell tests failed to reach significance. However, there was a tendency that students with MD used less frequently the simultaneous enumeration ($\chi^2(1) = 3.21$, $p = .073$), but more often the counting all process ($\chi^2(1) = 2.72$, $p = .098$). Figure 3 indicates the distribution of enumeration processes for each condition: subitizing range (random arrangement), counting range (random arrangement), and canonical arrangement. Within the subitizing range, 180 items were analyzed (2–4 in three different arrangements), within the counting range 300 items were analyzed (5–9 in three different arrangements), and in the canonical arrangement condition 160 items were analyzed (2–9 in one arrangement).

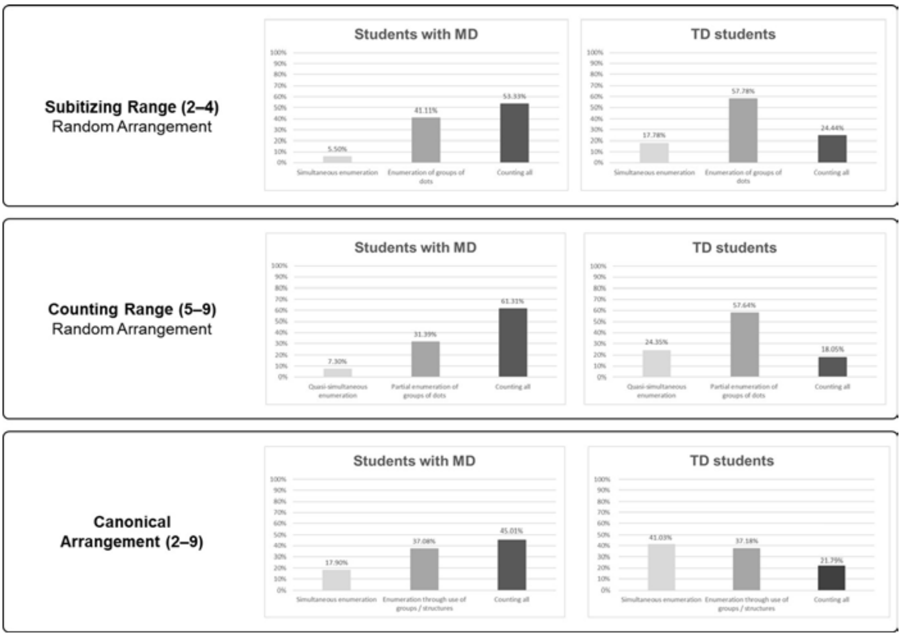


Figure 3. **Distribution of enumeration processes in three conditions**

Discussion

The aim of this explorative study was to investigate if students with and without MD differ in enumeration processes of small sets of objects (1–9) on group level. We studied error rates, response times, numbers of fixations, and enumeration processes based on qualitative ET video data of students with and without MD (*n*=10 each).

One obvious limitation of this study is the small sample size of *N*=20 students—with according limitations concerning the statistical analysis. Yet, our study using ET to investigate enumeration processes had an explorative nature. Similar studies using ET to investigate students with MD in quantity recognition had similar or even smaller numbers of participants (see, e.g., Moeller et al., 2009). Our kind of analysis, inferring enumeration processes from gaze-overlaid videos and working out categories of processes inductively, requires extensive data processing. Therefore, we think that our sample size was reasonable for an explorative study. Our inductive approach led to a (tentative) set of categories of enumeration processes of small sets of items. Such enumeration processes for small sets in random and canonical, dice-like arrangements were—to the best of our knowledge—not empirically found in previous studies on students with MD through the use of ET video data and, thus, constitute a novel finding. Yet, interpreting the results of our study, it is important to bear in mind that the study had little statistical power.

Enumeration: General results

We studied student enumeration in three conditions: First, we investigated student enumeration in the *subitizing range*, i.e., in random arrangements of sets of 1 up to 4 dots. Second, we investigated student enumeration in the *counting range*, i.e., in random arrangements of sets of 5 up to 9 dots. Third, we investigated student enumeration in *canonical arrangements* of 1 up to 9 dots that were arranged in dice-like patterns (Figure 1). In our study with grade 5 students, both groups of students—with and without MD—had low error rates in general: Students were generally well able to enumerate the presented dots in one way or another. This is not surprising, since young children are normally able to grasp sets of items as presented in our study (e.g., Starkey & Cooper, 1980, 1995). Our results indicate that students need more time to enumerate bigger quantities than smaller ones: Response times in the counting range were longer than in the subitizing range. This connects to previous findings, for example, by Schleifer and Landerl (2011). For canonical arrangements (1–9), response times were low as compared to the counting range, which connects to Ashkenazi et al.'s (2013) finding that response times are slower in the random arrangement than in the canonical arrangement.

Our ET study revealed a set of student enumeration processes for randomly arranged small sets in the subitizing and counting range as well as for canonical arrangements. For all three conditions, we were able to distinguish three kinds of enumeration processes: (1) The *(quasi-) simultaneous enumeration*, where participants grasped all dots at once (in subitizing range and canonical condition) or perceived them quasi-simultaneously (in counting range condition), (2) the (*partial) enumeration of groups*, where participants made use of patterning abilities partially, and counted the rest of the items, and (3) *counting all*, where students counted all items that were presented. To the best of our knowledge, analysis of gaze-overlaid videos has not been used to investigate enumeration processes and group differences in these processes for randomly and canonically arranged small sets before. We think that our work might be an impulse for further studies to investigate enumeration processes.

Enumeration of students with mathematical difficulties

The results of our study indicate that students with MD differ from TD students in enumeration. In *error rates*, however, such differences were hardly perceivable. Even though the error rates were slightly higher for the students with MD (4.7%) than for TD students (2.8%), group differences were not significant. This finding relates to Schleifer and Landerl (2011) who also found low error rates in students attending grades 2 to 4. We cannot confirm Ashkenazi et al.'s (2013) finding that students with MD made significantly more errors than their controls.

The *response times* tended to be longer for students with MD. Group differences were significant only for the canonical arrangements, indicating that students with MD were slower in recognizing and naming dice-like patterns than their TD peers. Yet, we cannot confirm results by Landerl (2013) and van der Sluis et al. (2004) who found significantly longer response times in students with MD also for random arrangements. It is interesting that we found group differences in response times particularly in the canonical arrangement: The discrepancy appears to be higher in the canonical arrangement than in subitizing or counting of randomly arranged items. This result supports Ashkenazi et al.'s (2013) finding that students with MD benefit less than their TD peers from the canonical arrangement and that they may have according "difficulties in implicit pattern recognition" (p. 43). Since students with and without MD in our study did not differ in visuo-spatial working memory performances, possible deficits in visuo-spatial working memory appeared not to be causal for lower performances of students with MD in the canonical arrangement. Since dice-like patterns are often learned in children's out-of-school lives, for instance, in board games, it is possible that a lack of such experiences led to the slower recognition of these patterns by students with MD, or that they benefited less from such experiences

Regarding *numbers of fixations*, we found that students with MD tended to have more fixations than TD students, which may relate to their longer response times: In the longer time span that they looked at the tasks, they had more fixations on the tasks. However, group differences in numbers of fixations were not significant. Thus, we cannot confirm the results of the case study by Moeller et al. (2009), which had shown more fixations for a student with MD than for students without MD.

As compared to the above-mentioned analyses of error rates, response times and numbers of fixations, the analysis of *enumeration processes*—based on the analysis of gaze-overlaid videos—revealed many differences between students with and without MD. We found significant differences in the distribution of student enumeration processes between students with and without MD in all conditions: in the subitizing range in random arrangements, in the counting range in random arrangements, and in the canonical arrangement. This is a strong result given the small number of participants. Looking closer at the three conditions, we found that in the *subitizing range*, the overall distribution of enumeration processes was significantly different between students with and without MD. On enumeration process level, we found that students with MD counted all dots more often than TD students. For the *counting range*, we also found significant differences of the overall distribution of enumeration processes between students with and without MD. In particular, we found that students with MD less often perceived dots quasi-simultaneously than their TD peers (through groupitizing) and less often used a partial enumeration of groups of dots, but that they counted all dots more often than TD students. For the *canonical arrangement*, the distribution of enumeration was likewise significantly different between the groups of students with and without MD. Here, we also found a tendency of students with MD to perceive dots simultaneously (recognizing the dice-pattern) less often, while counting all dots more often than their TD peers. Our results relate to Moeller et al.'s (2009) finding that two boys with MD counted, while the control group was able to enumerate quantities in parallel or rather simultaneously.

Implications for practical work with students with MD

It appears that students' enumeration processes of small sets have diagnostical potential for identifying students with MD. For example, in random arrangements of dots, students with MD counted all dots significantly more often, whereas they significantly less often used quasi-simultaneous enumeration—even at the age of almost eleven in our sample. For school teachers, it may be valuable to pay attention to these indicators for identifying students with MD who are in need of further support. We also found that students with MD less often used groups of dots for enumeration (groupitizing), which indicates a lesser use of the part-whole schema (e.g., Krajewski & Schneider, 2009)—a basic concept of number acquisition. Cognitive processes of enumeration involving the understanding of number composition develop during childhood and are linked to students' later mathematical development (Starkey, & McCandliss, 2014). Our study indicates that even after four years of primary school, students with MD still tend to differ in these basic processes. For school instruction and interventions for students with MD, it appears to be important to support students' understanding of number composition and part-whole schema—not only in interventions at an early age (like in Moser Opitz et al., 2018, where part-whole relationship was a topic in an intervention with second graders), but still in fifth grade. Further, the fact that students with MD recognize dice patterns slower than TD students and that they make use of different enumeration processes than TD students indicates that students with MD might have less experiences with dice games or may have benefited less from such earlier experiences, and that they might be in need of further learning (or playing) opportunities involving dice patterns at school or at home.

Conclusion and Outlook

In conclusion, the novel findings from our explorative study may give a tentative answer to van der Sluis et al.'s (2004) question of whether students with MD are "slower in subitizing, or whether they [rather use] … a safer counting strategy" (p. 260): Our findings suggest that students with MD use qualitatively different enumeration processes than their TD peers. The qualitative analysis of student enumeration processes based on eye tracking and gaze-overlaid videos was crucial for our findings on group differences, since the other measures hinted at group differences at best: error rates and numbers of fixations showed only trends of group differences, and group differences in response times were significant in just one condition (canonical arrangement). We understand our explorative study as springboard for further research on student enumeration processes in small set enumeration tasks and on differences in enumeration performances of students with and without MD. We think that studies with bigger samples will overturn the limitations of our study.

References

Ashkenazi, S., Mark-Zigdon, N., & Henik, A. (2013). Do subitizing deficits in developmental dyscalculia involve pattern recognition weakness? *Developmental Science*, *16*, 35–46. https://doi.org/10.1111/j.1467-7687.2012.01190.x

Bugden, S., & Ansari, D. (2016). Probing the nature of deficits in the 'approximate number system' in children with persistent developmental dyscalculia. *Developmental Science, 19*, 817–833. https://doi.org/10.1111/desc.12324

Clarke, B., Cheeseman, J., & Clarke, D. (2006). The mathematical knowledge and understanding young children bring to school. *Mathematics Education Research Journal*, *18*, 78–102. https://doi.org/10.1007/BF03217430

Clements, D. H. (1999). Subitizing: What is it? Why teach it? *Teaching Children Mathematics*, *5*, 400–405.

Cohen, J. (1988). *Statistical power analysis for the behavioral science.* Hillsdale, NJ: Lawrence Erlbaum.

Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* [Research methods and evaluation in the social and human sciences]. Berlin, Germany: Springer. https://doi.org/10.1007/978-3-642-41089-5

Fischer, B., Gebhardt, C., & Hartnegg, K. (2008). Subitizing and visual counting in children with problems in acquiring basic arithmetic skills. *Optometry & Vision Development, 39*, 24–29.

Gelman, R., & Gallistel, C. R. (1986). *The child's understanding of number.* Cambridge, MA: Harvard University Press.

Gray, S. A., & Reeve, R. A. (2014). Preschoolers' dot enumeration abilities are markers of their artihmetic competence. *PloS One, 9*, e99428.

Haffner, J., Baro, K., Parzer, P., & Resch, F. (2005). *Heidelberger Rechentest* [Heidelberger Calculation Test*]*. Göttingen, Germany: Hogrefe.

Hasselhorn, M., Schumann-Hengsteler, R., Gronauer, J., Grube, D., Mähler, C., Schmid, I., Seitz-Stein, K., & Zoelch, C. (2012). *Arbeitsgedächtnisbatterie für Kinder von 5 bis 12 Jahren* [Working memory battery for children from 5 to 12 years]. Göttingen, Germany: Hogrefe.

Krajewski, K., & Schneider, W. (2009). Early development of quantity to number-word linkage as a precursor of mathematical school achievement and mathematical difficulties: Findings from a four-year longitudinal study. *Learning and Instruction*, *19*, 513–526. https://doi.org/10.1016/j.learninstruc.2008.10.002

Kuhn, J.-T., Raddatz, J., Holling, H., & Dobel, C. (2013). Dyskalkulie vs. Rechenschwäche: Basisnumerische Verarbeitung in der Grundschule [Dyscalculia vs. severe math difficulties: Basic numerical capacities in elementary school]. *Lernen und Lernstörungen*, *2*, 229–247. https://doi.org/10.1024/2235-0977/a000044

Landerl, K. (2013). Development of numerical processing in children with typical and dyscalculic arithmetic skills—a longitudinal study. *Frontiers in Psychology*, *4*, 459. https://doi.org/10.3389/fpsyg.2013.00459

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174. https://doi.org/10.2307/2529310

Logie, R. H., & Pearson, D. G. (1997). The inner eye and the inner scribe of visuo-spatial working memory: Evidence from developmental fractionation. *European Journal of Cognitive Psychology*, *9*, 241–257. https://doi.org/10.1080/713752559

Mandler, G., & Shebo, B. J. (1982). Subitizing: an analysis of its component processes. *Journal of Experimental Psychology: General*, *111*, 1–22. https://doi.org/10.1037/0096-3445.111.1.1

Mayring, P. (2014). *Qualitative content analysis: theoretical foundation, basic procedures and software solution*. Klagenfurt, Austria: Beltz. https://www.ssoar.info/ssoar/handle/document/39517

Mazzocco, M. M., Feigenson, L., & Halberda, J. (2011). Impaired acuity of the approximate number system underlies mathematical learning disability (dyscalculia). *Child Development, 82*, 1224–1237. https://doi.org/10.1111/j.1467-8624.2011.01608.x

Mock, J., Huber, S., Klein, E., & Moeller, K. (2016). Insights into numerical cognition: Considering eye-fixations in number processing and arithmetic. *Psychological Research*, *80*, 334–359. https://doi.org/10.1007/s00426-015-0739-9

Moeller, K., Neuburger, S., Kaufmann, L., Landerl, K., & Nuerk, H. C. (2009). Basic number processing deficits in developmental dyscalculia: Evidence from eye tracking. *Cognitive Development*, *24*, 371–386. https://doi.org/10.1016/j.cogdev.2009.09.007

Moser Opitz, E., Freesemann, O., Prediger, S., Grob, U., Matull, I., & Hußmann, S. (2016). Remediation for students with mathematics difficulties: An intervention study in middle schools. *Journal of Learning Disabilities*, *50*, 724–736.

Moser Opitz, E., Grob, U., Wittich, C., Häsel-Weide, U., & Nührenbörger, M. (2018). Fostering the Computation Competence of Low Achievers through Cooperative Learning in Inclusive Classrooms: A Longitudinal Study. *Learning Disabilities: A Contemporary Journal*, *16*, 19–35.

Peterson, S. A., & Simon, T. J. (2000). Computational evidence for the subitizing phenomenon as an emergent property of human cognitive architecture. *Cognitive Science, 24*, 93–122. https://doi.org/10.1207/s15516709cog2401_3

Piazza, M., Mechelli, A., Butterworth, B., & Price, C. J. (2002). Are subitizing and counting implemented as separate or functionally overlapping processes?. *Neuroimage*, *15*, 435–446. https://doi.org/10.1006/nimg.2001.0980

Scherer, P., Beswick, K., DeBlois, L., Healy, L., & Opitz, E. M. (2016). Assistance of students with mathematical learning difficulties: how can research support practice? *ZDM Mathematics Education, 48*, 633–649. https://doi.org/10.1007/s11858-016-0800-1

Schindler, M., Bader, E., Lilienthal, A. J., Schindler, F., & Schabmann, A. (2019). Quantity recognition in structured whole number representations of students with mathematical difficulties: An eye-tracking study. *Learning Disabilities: A Contemporary Journal, 17*, 5–28.

Schindler, M., & Lilienthal, A. J. (2018). Eye-tracking for studying mathematical difficulties—also in inclusive settings. In E. Bergqvist, M. Österholm, C. Granberg, & L. Sumpter (Eds.), *Proceedings of the 42nd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 115–122). Umeå, Sweden: PME

Schleifer, P., & Landerl, K. (2011). Subitizing and counting in typical and atypical development. *Developmental Science*, *14*, 280–291. https://doi.org/10.1111/j.1467-7687.2010.00976.x

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*, 591–611. https://doi.org/10.1093/biomet/52.3-4.591

Simon, H, & Grünke, M. (2010). *Förderung bei Rechenschwäche* [Support for dyscalculia]. Stuttgart, Germany: Kohlhammer.

Starkey, P., & Cooper, R. G. (1980). Perception of numbers by human infants. *Science, 210*, 1033–1035. https://doi.org/10.1126/science.7434014

Starkey, P., & Cooper, R. G. (1995). The development of subitizing in young children. *British Journal of Developmental Psychology, 13*, 399–420. https://doi.org/10.1111/j.2044-835X.1995.tb00688.x

Starkey, G. S., & McCandliss, B. D. (2014). The emergence of "groupitizing" in children's numerical cognition. *Journal of Experimental Child Psychology, 126*, 120–137. https://doi.org/10.1016/j.jecp.2014.03.006

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

Van Viersen, S., Slot, E. M., Kroesbergen, E. H., Van't Noordende, J. E., & Leseman, P. P. (2013). The added value of eye-tracking in diagnosing dyscalculia: A case study. *Frontiers in Psychology*, *4*, 679. https://doi.org/10.3389/fpsyg.2013.00679

Von Glasersfeld, E. (1982). Subitizing: The role of figural patterns in the development of numerical concepts. *Archives de Psychologie, 50*, 191–218.

Wartha, S., & Schulz, A. (2012). *Rechenproblemen vorbeugen. Grundvorstellungen aufbauen. Zahlen und Rechnen bis 100* [Prevent calculation problems. Build up basic ideas. Numbers and arithmetic up to 100]. Berlin, Germany: Cornelsen.

World Health Organisation (WHO) (Ed.) (2019). International statistical classification of diseases and related health problems. 10th Revision. https://icd.who.int/browse10/2019/en Accessed 2020-10-12.