Small number enumeration processes of deaf or hard-of-hearing students: A study using eye tracking and artificial intelligence

Maike Schindler[1]*, Jan H. Doderer[1], Anna L. Simon[1], Erik Schaffernicht[2], Achim J. Lilienthal[2] and Karolin Schäfer[1]

[1]Department of Special Education and Rehabilitation, Faculty of Human Sciences, University of Cologne, Cologne, Germany, [2]School of Science and Technology, Örebro University, Örebro, Sweden

Students who are deaf or hard-of-hearing (DHH) often show significant difficulties in learning mathematics. Previous studies have reported that students who are DHH lag several years behind in their mathematical development compared to hearing students. As possible reasons, limited learning opportunities due to a lesser incidental exposure to numerical ideas, delays in language and speech development, and further idiosyncratic difficulties of students who are DHH are discussed; however, early mathematical skills and their role in mathematical difficulties of students who are DHH are not explored sufficiently. In this study, we investigate whether students who are DHH differ from hearing students in their ability to enumerate small sets (1−9)—an ability that is associated with mathematical difficulties and their emergence. Based on a study with $N = 63$ who are DHH and $N = 164$ hearing students from third to fifth grade attempting 36 tasks, we used eye tracking, the recording of students' eye movements, to qualitatively investigate student enumeration processes. To reduce the effort of qualitative analysis of around 8,000 student enumeration processes (227 students x 36 tasks), we used Artificial Intelligence, in particular, a clustering algorithm, to identify student enumeration processes from the heatmaps of student gaze distributions. Based on the clustering, we found that gaze distributions of students who are DHH and students with normal hearing differed significantly on a group level, indicating differences in enumeration processes, with students who are DHH using advantageous processes (e.g., enumeration "at a glance") more often than hearing students. The results indicate that students who are DHH do not lag behind in small number enumeration as compared to hearing students but, rather, appear to perform better than their hearing peers in small number enumeration processes, as well as when conceptual knowledge about the part-whole relationship is involved. Our study suggests that the mathematical difficulties of students who are DHH are not related to difficulties in the small number enumeration, which offers interesting perspectives for further research.

Introduction

Studies throughout recent decades repeatedly indicate that students who are deaf or hard-of-hearing (DHH) tend to have mathematical difficulties, which appear to be severe in many cases (e.g., Traxler, 2000; Blatto-Vallee et al., 2007; Qi and Mitchell, 2012; Marschark et al., 2013; Pagliaro, 2015). Even at preschool age, most students who are DHH show difficulties in their mathematical development (e.g., Kritzer, 2009; Pagliaro and Kritzer, 2013), and these difficulties tend to increase through school age and into adulthood (Bull et al., 2011; Gottardis et al., 2011).

Students with hearing loss can be referred to as DHH, depending on their degree of hearing loss or the hearing threshold (World Health Organization, 2021). A distinction is made between mild, moderate, moderately severe, severe, profound, or complete hearing loss (World Health Organization, 2021). In many cases, both ears are affected. Hearing loss might lead to difficulties in hearing auditory stimuli and communicating with others (Spencer and Marschark, 2010). Moreover, congenital hearing loss has an impact on learning and social development (Knoors and Marschark, 2014). With increased early detection and quality of hearing technology and early intervention, today, many children who were born congenitally deaf can communicate in spoken language due to their hearing technology (Spencer and Marschark, 2010). The increasing importance of bimodal-bilingual education has led to many children today having the opportunity to acquire both spoken and sign language from an early age (Marschark and Knoors, 2012). As a trend, the proportion of students receiving mainstream education has increased in recent years in many countries, and, at the same time, sign language remains an important part of the diverse DHH education (Leigh and Marschark, 2016; Marschark and Leigh, 2016). Despite this, school education for students who are DHH widely varies internationally. Didactical concepts may also differ depending on hearing status, communication modality, and additional needs, both in special schools and in mainstream education.

Mathematical abilities of students who are DHH have received relatively little attention as compared to, for example, their language and literacy skills, which is possibly related to "a pervasive belief that mathematics is 'not as important' as language and/or literacy" (Pagliaro and Kritzer, 2013, p. 150) among practitioners and researchers working with students who are DHH. Yet, the scarce body of literature on the mathematical abilities of students who are DHH has consistently indicated similar trends: That, they, on average, appear to lag in their mathematical development by several years as compared to their hearing peers. In the 1960s, 1970s, and 1980s, different studies suggested that students who are DHH perform significantly below their same-aged peers with normal hearing (e.g., Wollman, 1965) and that their mathematical development is delayed by 2 to 5 years (Hine, 1970; Wood et al., 1983). These

results were validated in studies within the last 30 years, which confirmed that students who are DHH, despite the improvement of early diagnostics, support, hearing technology, educational opportunities, and a greater appreciation of sign language, show a developmental delay of about 3 to 4 years in their mathematical abilities at primary and secondary school levels as compared to their hearing peers (e.g., Heiling, 1995; Frostad, 1996; Traxler, 2000; Qi and Mitchell, 2012; Edwards et al., 2013; Pagliaro, 2015). This discrepancy appears to increase until the age of 16 (Gottardis et al., 2011) and persist even in adulthood (Bull et al., 2011). Taken together, previous studies indicate that students who are DHH tend to have mathematical difficulties that often start as early as preschool age, increase with age, and persist in adulthood. These difficulties apply to a multitude of mathematical topics, including counting (e.g., Pagliaro and Kritzer, 2013), measurement (e.g., Austin, 1975; Pagliaro and Kritzer, 2013), fractions (e.g., Titus, 1995; Bull, 2008; Mousley and Kelly, 2018), number line estimation (e.g., Bull et al., 2011; Bedoya-Ríos and Dorneles, 2021), and word problems (e.g., Kelly and Mousley, 2001; Nunes and Moreno, 2002; Hyde et al., 2003; Ansell and Pagliaro, 2006; Blatto-Vallee et al., 2007). One of the relative strengths of students who are DHH lies in the area of geometry (Pagliaro and Kritzer, 2013), which is presumably related to strong visual and spatial skills, wherein students who are DHH are assumed to have an advantage (e.g., visuospatial working memory, Marschark and Knoors, 2012). For example, children who are DHH were found to be able to remember complex figures better than hearing students, yet, if the figures move and the sequences of movements are then to be described and classified in temporal order, students who are DHH show an inferior performance compared to hearing students (Todman and Seedhouse, 1994; Marschark and Knoors, 2012).

Mathematical delays of students who are DHH appear to start before schooling. When school begins, there seems to be a quantitative delay of about 2 years in mathematical abilities of children who are DHH as compared to hearing children (Swanwick et al., 2005; Pagliaro and Kritzer, 2013). Studies by Kritzer (2009) and Pagliaro and Kritzer (2013) have furthermore indicated that even at preschool age a majority of children who are DHH showed delays in early mathematics, for example, in number comparisons, counting by groups, and word problems (Kritzer, 2009; Pagliaro and Kritzer, 2013). Leybaert and Van Custem (2002) found in preschoolers that "deaf children exhibited age-related lags in their knowledge of the number sequence; they made different errors from those of hearing children, reflecting the rule-bound nature of sign language. Remarkably, their performance in object counting and creating sets of given cardinalities was similar to that of hearing children" (p. 482). Furthermore, Zarfaty et al. (2004) found that preschool children who are deaf and using sign language did not have difficulties in representing and discriminating numbers and that they performed better in spatial (not temporal) tasks

than their hearing peers. Yet, it is important to consider that many studies investigated deaf native sign language users and their understanding of number concepts. Even though it is not entirely clear how children who are DHH with hearing parents use sign language, or how children who are DHH use spoken language perform, it was found that better sign language skills of students who are DHH tend to result in better mathematical performance (e.g., Henner et al., 2021). Taken together, the results on the mathematics performance of preschoolers are diverse, indicating that it is not quite easy to say where the mathematical difficulties of students who are DHH originate, yet certain delays appear to be present even before the school age.

What are the reasons for the difficulties in mathematics of students who are DHH, which appear to start in preschool? One obvious reason may be a delay in language development, which affects many children who are DHH (Spencer and Marschark, 2010). The development of language is closely linked to the development of mathematical thinking and is acquired incidentally and intuitively in everyday life (Nunes and Moreno, 2002). Social activities, such as counting, are learned and taught through language. In this context, children's access to mathematical knowledge is initially informal, without the content being explicitly labeled as mathematical. For example, many nursery rhymes contain counting activities, shape comparisons, or spatial descriptions that are being automated through frequent singing and, later, make it easier for children to acquire counting and mathematical problem-solving skills. Several other factors have been suggested to explain why students who are DHH lag in their mathematical development, which can be broadly divided into two categories: One possible explanation is that the reason lies in fewer learning opportunities for DHH for students that are related to limited incidental exposure to numerical ideas, language issues, and barriers, parents making few references to mathematical concepts in everyday activities, and quality of education (Titus, 1995; Nunes, 2004; Kritzer, 2008, 2009; Barbosa, 2014); the other explanation assumes that the problems originate from students who are DHH themselves, and idiosyncratic difficulties of students who are DHH in basic numerical processing and retrieval processes (e.g., for a detailed discussion, see Zarfaty et al., 2004; Bull et al., 2006; Pagliaro and Kritzer, 2013). In a review of the literature, Santos and Cordes (2022) found that in addition to limited or reduced language access, especially in young children, factors, such as executive functions, specifically, working memory, may also be decisive for the fact that students who are DHH may have difficulties in mathematics. Some previous studies have investigated the basic numerical processing of students who are DHH, in particular. For example, Bull et al. (2006) found that university students who are DHH did not differ from their hearing peers in response times in small-number enumeration processes, indicating that these processes appear not to differ for students who are DHH at

the university level. However, the participants in this study were at a high academic level. It is not yet clear whether small number enumeration processes of school-age students who are DHH differ from those of their hearing peers and whether this may be one of the causes for their mathematical difficulties, or whether the difficulties, instead, originate from other factors.

Mathematical difficulties in hearing children are often associated with issues in basic numerical processing (e.g., Butterworth, 2005; Wilson and Dehaene, 2007), and such basic numerical difficulties tend to cascade to severe difficulties in a wide range of mathematical abilities (Butterworth, 2005; Geary, 2013). One of such abilities known to predict mathematical achievement in later school years is a small-number enumeration, the ability to grasp sets of items and say how many there are (Starkey and Cooper, 1995). For children at preschool level and at the beginning of primary school, it is crucial to learn to enumerate quantities (that is, to perceive sets of items and say how many there are) (e.g., Department of Education, 2013). The ability to enumerate small sets of items involves different processes: (perceptual) subitizing, counting, and conceptual subitizing, so-called groupitizing (Gelman and Gallistel, 1986; Schleifer and Landerl, 2011; Ashkenazi et al., 2013; Starkey and McCandliss, 2014). Subitizing means the ability to enumerate small quantities without counting in a fast and exact way; the process is perceptual, automatized, and often subconscious (Clements, 1999; Fischer et al., 2008). Humans—even young children—are usually capable of perceiving quantities of up to four items through subitizing (Starkey and Cooper, 1980, 1995; Mandler and Shebo, 1982; Clarke et al., 2006; Schleifer and Landerl, 2011). Subitizing is a foundation for the development of the concept of numbers and arithmetic learning (Starkey and Cooper, 1995). Humans cannot (in most cases) subitize sets of more than four items; larger numbers of items need to be perceived serially, for example, through counting or structuring them in groups (that can again be subitized). Counting is a process that is socially taught, and it involves conventions that have been summarized by Gelman and Gallistel (1986) as counting principles: *the one-to-one principle*, meaning that each item is assigned exactly with one number word, *the stable-order-principle*, meaning that the order of number words is predetermined (1, 2, 3, 4, …), *the cardinal principle*, meaning that the number word that was said last indicates the cardinality of the set, *the abstraction principle*, meaning that different kinds of objects can be counted, irrespective of their appearance, and *the order-irrelevance principle*, meaning that the order, in which the items are counted, is irrelevant for the cardinality of the set. Compared to subitizing, counting is a "slower and more error-prone process of one-to-one mapping between a set of objects and number words" (Schleifer and Landerl, 2011, p. 280). In the context of studies addressing enumeration processes, researchers refer to the counting range when sets of

five or more items are to be enumerated, and to the subitizing range for sets of one to four items, even though, of course, counting processes can also be involved in the enumeration of sets within the subitizing range (e.g., Schleifer and Landerl, 2011; Ashkenazi et al., 2013; Schindler et al., 2020b). Despite subitizing and counting processes, in enumeration, processes of perceiving sets in subsets, dividing sets, and combining subsets to a set are involved (Clements, 1999). This involves the part-whole schema (Starkey and McCandliss, 2014), which means "understanding that quantities can be decomposed into pieces and reassembled again" (Krajewski and Schneider, 2009, p. 513). The enumeration of sets in subsets (groups) is called groupitizing (Starkey and McCandliss, 2014) or "conceptual subitizing" (Clements, 1999) since conceptual processes of the part-whole relationship are involved. Groupitizing allows sets of items to be perceived quasi-simultaneously, that is, by subitizing subsets (Anobile et al., 2020; Wege et al., 2022). Furthermore, enumeration of small sets can involve the recall of familiar patterns, such as dice patterns. In this case, researchers talk about canonical representations (e.g., Ashkenazi et al., 2013). Such patterns do not necessarily have to be counted, but they can be recalled if the image of the pattern is familiar, and, when recognized, the according number can be recalled (e.g., Starkey and McCandliss, 2014). In this case, the child does not need to perceive items serially but can associate the pattern to a number word (Von Glasersfeld, 1982). In general, the canonical representations are symmetrical and, thus, allow the use of patterns or symmetries for number enumeration (Hsin et al., 2021).

Enumeration of small sets is essential for learning basic arithmetic and, therefore, crucial to the development of students' mathematical skills (Starkey and Cooper, 1995). For example, preschoolers' mastery of structured sets (e.g., dice patterns) can predict their arithmetic abilities in grade 1 (Kreilinger et al., 2021), and their subitizing ability can predict their arithmetic performance in later school years (Hannula-Sormunen et al., 2015). Mathematical difficulties (MD)—that is, difficulties in understanding basic arithmetic concepts (e.g., Scherer et al., 2016; Moser Opitz et al., 2017)—are associated (among other factors) with deficits in subitizing that affect enumeration processes (e.g., Butterworth, 2005; Wilson and Dehaene, 2007). Previous research on students' enumeration processes indicates a "dysfunctional subitizing mechanism" (e.g., Schleifer and Landerl, 2011, p. 280) for students with MD (Van der Sluis et al., 2004; Schleifer and Landerl, 2011; Landerl, 2013). Studies addressing, among other things, the students' response times in the enumeration of small sets found that students with MD (aged 7–17, Fischer et al., 2008; aged 10, Moeller et al., 2009) were slower than students without MD in the subitizing range. When examining the enumeration abilities of students in second to fourth grade, students with MD were found to have subitizing problems, specifically, steeper response times slopes (Schleifer and Landerl, 2011; Landerl, 2013). Response

time slopes in the counting range were similar for students with and without MD (Schleifer and Landerl, 2011). Further, Van der Sluis et al. (2004) found that fourth to fifth graders with low mathematical skills needed more time for enumeration of sets in the subitizing range. Gray and Reeve (2014) found that weak subitizing profiles of preschoolers were related to poor arithmetic (addition) skills, providing another indicator of difficulties in subitizing for children with MD in their study. Ashkenazi et al. (2013), in a study distinguishing the arrangement of items (canonical vs. random), found that students with MD had higher error rates than the control group when enumerating canonically arranged items, which increased as the number of items increased. For canonically arranged item patterns, longer response times in a group of students with MD—as compared to a control group—were also found by Schindler et al. (2020b). This suggests that students with MD benefit less from canonical arrangements of items (Ashkenazi et al., 2013; Schindler et al., 2020b). Based on product-related data, such as response times or error rates, causes of differences between performances of children with and without MD cannot be definitively explained (Van der Sluis et al., 2004).

However, deficits in subitizing, which may be involved in the emergence of MD, can be further investigated by the analysis of eye movements during enumeration processes (see Mock et al., 2016). To find out whether students with MD are only slower at subitizing or whether they rely on qualitatively different enumeration processes, eye-tracking studies (Moeller et al., 2009; Schindler et al., 2020b) have proven to provide insights into students' enumeration processes. In a case study with students of the same age as the students in this study, Moeller et al. (2009) investigated enumeration processes for small sets of 1 to 8 items. They found differences between children with and without MD, in that counting processes were observed in both students with MD. Another ET study (N=20) by Schindler et al. (2020b) also showed differences in enumeration processes between students with and without MD. In both the subitizing (2–4) and the counting (5–9) range, processes of counting all items were observed more frequently for students with MD as compared to their peers without MD, who used (quasi-)simultaneous enumeration more often (Schindler et al., 2020b). These results were also found for the enumeration of items in canonical arrangements (2–9). In summary, qualitatively different processes between students with and without MD were indicated.

This article intends to contribute to the investigation of possible roots of mathematical difficulties of students who are DHH. Specifically, we want to explore if the mathematical difficulties of students who are DHH go along with deficits in small-number enumeration. This study aims to investigate if students who are DHH differ from hearing students in small-number enumeration processes. We ask the research

question: "*Do DHH students differ from hearing students in small number enumeration processes?*" To investigate enumeration processes, we conducted a study using eye tracking (ET), the recording of students' eye movements, since ET has shown to be a valuable tool for investigating small number enumeration processes (Moeller et al., 2009; Schindler et al., 2020b). In a study with $N = 227$ students (164 hearing students, 63 students who were DHH), we analyzed eye movements of approximately 8,000 students' enumeration processes (227 students x 36 tasks) and—to reduce the effort of qualitative analysis in a large-scale study with several thousand items—combined it with Artificial Intelligence (AI) to identify student enumeration processes from the heatmaps of student gaze distributions.

With the focus on basic numerical difficulties as a possible root of MD, this paper addresses an issue that is not only of concern for researchers and educators in the domain of students who are DHH and their teaching but beyond: Addressing enumeration, specifically, counting principles or the part-whole schema early on is of concern for many students—for students with mathematical difficulties and students in the domain of special education (e.g., Krajewski and Schneider, 2009; Hecht et al., 2011; Schleifer and Landerl, 2011; Garrote et al., 2015). The diagnostics and support of students with difficulties in these aspects are of paramount importance since longitudinal studies indicate that students who enter school with low early-math skills often do not overcome their deficits during primary school (e.g., Viesel-Nordmeyer et al., 2019). Our study contributes to the research body on student enumerations processes and can be a springboard for further research on students with MD and/or special educational needs.

Materials and methods

Participants

A total of $N = 227$ students participated in this study (see Table 1). The sample consisted of a group of hearing students ($n = 164$) and a group of students who are DHH ($n = 63$). The hearing group consisted of students in fifth grade at a German comprehensive school in North Rhine-Westphalia (NRW) aged 9.10 to 12.6 years (mean: 10.9 years, SD: 0.7 years). The group of students who are DHH consisted of fifth-, fourth-, and third-grade students between 9.1 and 13.6 years (mean: 11.9 years, SD: 1) from German special schools for students who are DHH in NRW. The inclusion of students from third and fourth grades in the DHH group served the purpose of approximating the mean age of the participants in the two groups. This appeared to be beneficial, since most of the students who are DHH were in the same grade level (fifth grade) but in a higher school year chronologically than the hearing students. Furthermore, it allowed for a larger sample to be used.

TABLE 1 Participants in the study.

	Hearing group ($n = 164$)	DHH group ($n = 63$)
Participant information		
Age: mean (standard deviation)	10.9 (0.7)	11.9 (1.0)
Gender: n_{girls} (%$_{girls}$)	72 (43.91)	33 (52.38)
Grade (s): grade (%)	Grade 5 (100)	Grade 3 (6.35)
		Grade 4 (33.33)
		Grade 5 (60.32)
Mathematical abilities		
HRT Mean t-score (standard deviation)	40.37 (9.90)	35.04 (9.90)
Mathematical difficulties: n (%)	69 (42.07)	41 (65.08)
At risk zone: n (%)	36 (21.95)	9 (14.29)
Typically developing: n (%)	59 (35.98)	13 (20.63)

The degree of hearing loss in the group of children who are DHH varied, including mild ($n = 1$; 1.6%), moderate ($n = 19$; 30.2%), severe ($n = 20$; 31.7%), and profound ($n = 21$; 33.3%). Two students had a central auditory processing disorder (CAPD) (3.2%), and two other children were affected by single-sided deafness (SSD) (3.2%). Half of the students were fitted with bilateral hearing aids (32; 50.8%), and about a quarter had bilateral cochlear implants (CI) (15; 23.8%). Six students had a bimodal fitting with a hearing aid and a cochlear implant (9.5%). Three students were fitted with a unilateral cochlear implant (4.8%) and one had a unilateral hearing aid (1.5%). One student was fitted with bilateral bone-anchored hearing aids (BAHA) (1.5%). Five of the students, including the two with CAPD, were unaided. As for the students' familial first language (L1), 30 of the students had German as their first language, ten students were bilingual with German and another language (Arabic, Farsi, Italian, Russian, Serbian, Turkish, and Ukrainian), five had Turkish as first language, three Arabic, three German Sign Language (DGS), three Kurdish, and each one Afghan, Greek, Hungarian, Polish, Romanian, Syrian, Tamil, and Urdu, and one that is not specified. While conducting our study, we asked the students for their individual preferred communication modality: 28 of the students preferred spoken language as a communication modality (44.4%), 30 preferred sign-supported speech (47.6%), while five students preferred sign language (7.9%). Sign language, in this case, DGS, is based on a visual-spatial modality and has its own grammatical structure and syntax. While sign language is a language of its own, sign-supported speech only makes use of some of the DGS' signs to support spoken language while speaking. The diversity of the participants who are DHH in this study—especially about prior language experience, communication modality, hearing status, and hearing care—represents well the heterogeneity often encountered in special education schools for children who are DHH in Germany. We chose to include students who

attended schools for students who are DHH, who each had a special educational need in hearing, with the commonality of all children in the sample having a peripheral hearing loss and, in a few cases, a central auditory processing disorder, and we refrained from deliberately selecting children that fulfill a specific language or modality profile.

Before the study, legal guardians were informed about the aim and the procedures of the study and that students' participation was voluntary. The participating students were then informed individually about the content and procedures of the study; they had the opportunity to ask questions and to stop the study at their convenience. All students could choose what modality they preferred during the study (spoken language, sign-supported speech, or sign language). The interviewer, who conducted the study, was fluent in all three modalities. It was ensured that all students understood the tasks and felt comfortable during the study.

Prior to the ET study, a standardized arithmetic test, HRT (Haffner et al., 2005), was administered. The HRT is widely used in German-speaking countries to diagnose MD, and it has been used in previous studies examining small number enumeration (e.g., Moeller et al., 2009; Schleifer and Landerl, 2011; Schindler et al., 2019a, 2020b). HRT uses the percentile rank (PR) to diagnose MD, which provides information on what percentage of the norm sample performed equally well or worse (Haffner et al., 2005). If the PR is ≤10, the child is considered to have MD. Between PR 11 to 25, students are within an "at-risk zone" for MD. Students with PR>25 are considered typically developing. In this study, only the first half of the HRT, which focuses on numbers and arithmetic, was administered to the students (similar to Schleifer and Landerl, 2011; Landerl, 2013; Schindler et al., 2019a, 2020b). This arithmetic part alone can be used to diagnose MD (Haffner et al., 2005). The HRT is standardized from the end of first grade through the first six weeks of fifth grade. The sample of students with DHH in this study was tested toward the end of the school year. To use the HRT for fifth graders with DHH at the end of the school year, we adjusted the t-scores following a study by Pitters (2018) that also took place toward the end of the fifth-grade year. Following the average increase in t-scores of 5.5 from the beginning of grade 5 to the end of grade 5 found in Pitters' study, we decreased the t-scores of DHH fifth graders by 5.5 to adjust the values with respect to the standardization at the beginning of fifth grade.

Eye tracking as a research method in mathematics education

The human eye provides high resolution only in the small region of the fovea, the *fovea centralis* (Holmqvist et al., 2011). This region corresponds to a very small part of our field of view (only about the size of a thumbnail at arms' length).

Hence, the eyes need to move, so that the foveal area is oriented toward those regions in the environment from which detailed information is needed (Rayner, 1998; Henderson, 2003). Conversely, following the sequence of movements of the eye allows insight into visual attention and what cognitive processes were associated with the need to collect high-resolution visual information in the observed gaze pattern. Thus, ET, the recording of eye movements that brings areas of interest into the foveal vision, is used as a tool to study cognitive processes in many areas of research, including mathematics education research. ET investigates overt attention since people pay overt attention to those areas that they perceive by the foveal vision: When the eyes move directly to the stimulus, their attention is there, whereas those areas that are perceived peripherally may be taken in by covert attention (Posner, 1980; Carrasco, 2011).

ET is of growing interest in research in mathematics education (and, prospectively, also for educators) as a comparatively unobtrusive tool that allows access even to unconscious cognitive processes (Lilienthal and Schindler, 2019; Strohmaier et al., 2020; Schindler, 2021). Interest in ET has increased since ET devices became more affordable, easy to use, advanced, and accurate (Lilienthal and Schindler, 2019). Another reason for the increasing popularity of ET is that the inherent difficulty to draw conclusive inferences from gaze tracks, as well as the difficulty to extract useful information from a large amount of data, are mitigated by theoretical advances in interpretation (Schindler and Lilienthal, 2019). The interpretation of ET data is often based on the so-called eye-mind hypothesis (Just and Carpenter, 1980). This hypothesis states that there is no significant difference between what the eyes fixate on and what is processed (Holmqvist et al., 2011). Even though this hypothesis, which was developed in reading research (Just and Carpenter, 1980), does not generally hold true in mathematics (see Schindler and Lilienthal, 2019), previous research has shown that the interpretation of eye-tracking data in the domain of enumeration and quantity recognition is more straightforward (Schindler and Lilienthal, 2018; Schindler et al., 2019a, 2020b). Another reason for the growing interest in ET in mathematics education research is that computational resources for automated analysis of gaze data are available at a low cost. This and the other factors mentioned above have opened the prospect of novel applications of ET in mathematics education that use AI for automated or partially automated analysis of gaze data (Schindler et al., 2020a). During the last decade, mathematics education research has seen a notable increase in the number of publications that use ET (Lilienthal and Schindler, 2019; Strohmaier et al., 2020). An extensive review paper by Strohmaier et al. (2020) identified a large variety of topics, including studies of numbers and arithmetic (the largest group), learning difficulties, computer-supported learning, and studies of affective variables (Strohmaier et al., 2020).

Eye-tracking device and stimuli

We recorded eye movements with the screen-based eye tracker, Tobii Pro X3-120. This eye tracker allows the tracking of eye movements (binocular) with a sampling rate of 120 Hz. It is an eye-tracker shaped like a bar that can be mounted to the bottom of a screen's frame. Therefore, it is unobtrusive and not distracting to participants. Tobii Pro X3-120 uses infrared illuminators to create corneal reflection patterns from participants' eyes. According to the manufacturer, the accuracy is $0.4°$ (Tobii, 2019). The average accuracy in this study was $0.9°$. Participants can move during the recording without affecting accuracy and precision. A 9-point calibration was performed before each data collection.

In total, there were 36 items representing the sets from 1 to 9. Every quantity was presented in four different arrangements: The sets were each presented in a canonical arrangement and in three random arrangements, arranged differently each time (see Figure 1 for examples). For the random arrangements, two ranges can be distinguished: the subitizing range (2–4) and the counting range (5–9). The canonical arrangements were divided into the dice range (2–6) and beyond the dice range (7–9). The items were the same as in the study by Schindler et al. (2019b).

The enumeration tasks were presented on a 24" full HD screen with a refresh rate of 60 Hz and a resolution of 1,920 x 1,080 pixels. The dots on this screen had a diameter of 2 cm and a maximum span of the dots (vertical/horizontal) of 15 cm. The students sat on a chair about 60 cm from the monitor.

Machine learning to support the analysis of eye-tracking data

Our study uses ET in combination with machine learning (ML) to investigate student enumeration processes. Generally, the term ML on the one hand refers to a *subject of study* in AI: the study of computer algorithms that can improve automatically through experience and by using training data (Mitchell, 1997). In addition, ML refers to a specific *set of methods* "that can automatically detect patterns in data, and then use the uncovered patterns to predict future data, or to perform other kinds of decision making under uncertainty" (Murphy, 2012, p. 1). It is important to emphasize that often, no explicit manual coding is required. This goes along with one reason why ML is appealing: It promises to shorten or eliminate the laborious process of developing complex software. More importantly, ML allows learning models from examples, which represent patterns or dependencies that are otherwise unknown and cannot be coded analytically.

Among the three major types of ML, supervised learning, unsupervised learning, and reinforcement learning, only supervised and unsupervised machine learning have been used

so far in mathematics education, to the best of our knowledge. In this work, we use unsupervised learning. For a better understanding, we describe the general differences between supervised learning and unsupervised learning in this section and discuss related work that used unsupervised learning in mathematics education in the following. In predictive or supervised learning, the goal is to learn a mapping from inputs x to outputs y, given a labeled training set. After training, the learned mapping can be used to make categorical or nominal predictions (Murphy, 2012). Supervised learning (SL) was used, for example, in Schindler et al. (2019b), where the inputs **x** are (as in this paper) the heatmaps that represent the non-temporal information in ET sequences and the labels for each heatmap specify whether it belongs to a student with MD or a student who is typically developing. After training as described in Schindler et al. (2019b), the SL algorithm can classify unseen heatmaps and can be used to predict whether the corresponding student is likely to have MD or not. The second type of ML that has been previously used in mathematics education research is the descriptive or unsupervised learning approach (USL), where training samples **x** but no labels y are given. The computer is then tasked to "find 'interesting patterns' in the data" (Murphy, 2012, p. 2). This is also called "exploratory data analysis" or "knowledge discovery." As Murphy (2012) notes, USL is a much less well-defined problem than SL. This is because there is no a-priori guidance about what kinds of patterns to look for, and there is no obvious error metric to use, unlike SL, where a prediction y for a given **x** can be compared to the observed value (Xu and Wunsch, 2009). USL algorithms typically separate the unlabeled training data {**x**} into a number of meaningful clusters, and the major interest in USL is in such clustering algorithms. Clustering can be useful for several reasons, such as to get a compressed representation of the data or to generate hypotheses through data exploration (Aldenderfer and Blashfield, 1984).

In the context of mathematics education, Schindler et al. (2020a) have explored the possibility of identifying student strategies in whole number representations using ET combined with USL. Quantity recognition tasks on the 100-dot field were presented and, as in this paper, heatmaps were used as representations of gaze sequences. The result of a specific clustering algorithm, Self-Organizing Maps (SOMs, Kohonen, 2001), was then evaluated to find out whether the clusters were consistent with respect to enumeration strategies. The initial results were mixed. The clusters found with the SOM algorithm were consistent to some extent but sometimes also subsumed different strategies in one cluster. This result is not surprising as the clustering used only the visual similarity of heatmaps and discarded, for example, temporal information and information about absolute durations. Other clustering algorithms (for example, with tailored proximity metrics and an optimized intermediate representation) may improve well the consistency. However, an important aspect of Schindler et al. (2020a) study is that clustering was used to provide an independent, non-human

FIGURE 1
Examples of the items.
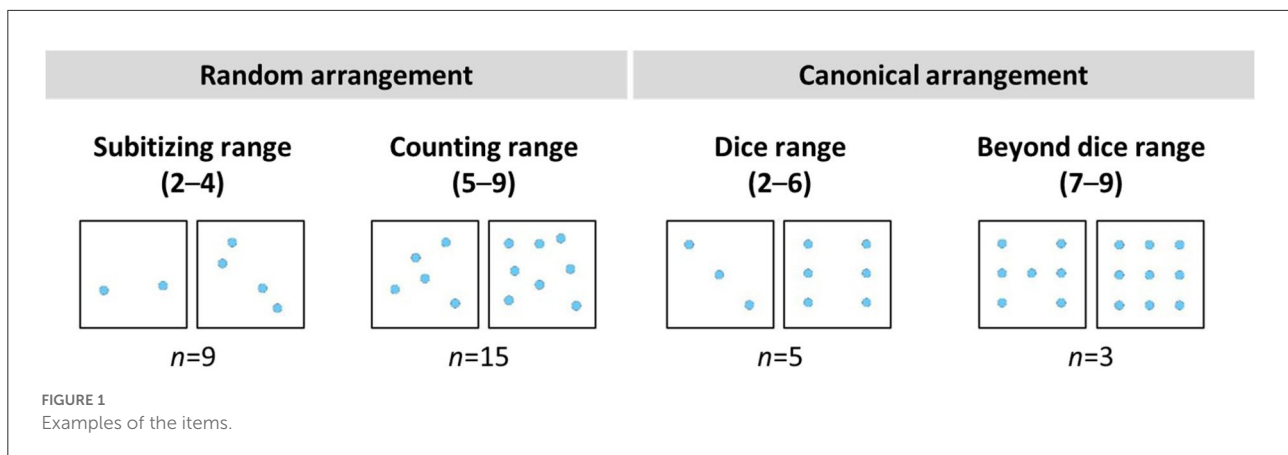
view of the data. In this sense, clustering can be seen as an example, in which AI is used to support human researchers—similar to a researcher colleague who takes an independent look at the data with a different perspective. In the future, we can imagine that different USL components, based on different clustering algorithms, generate different category hypotheses and, like an "AI colleague," make suggestions about meaningful ways to categorize the data. Human researchers would then interpret and verify these suggestions based on pre-studies with smaller numbers of participants and a principled understanding of the applied USL algorithms.

Procedure

The study took place in individual sessions in a quiet room at the respective schools of the students. Tasks were presented on the full HD screen with the eye tracker attached to the bottom frame. The study began with the calibration procedure, which was followed by three practice tasks to verify and ensure task comprehension. This was followed by 36 enumeration tasks. Participants were instructed to correctly name the number of dots represented as quickly as possible in their preferred communication modality (spoken language, sign-supported speech, or sign language). A fixation star was displayed between tasks in the center of a white screen to separate each task: Students were instructed to fixate on the star as long as it was visible so that students' gazes in the tasks always started from the very same position, the center of the screen. The items were intermixed randomly, using the same order for each student. The students were not given feedback on their responses. Students' responses were recorded with an audio recorder to analyze if they answered correctly (for example, "six" for six dots) or incorrectly. For students who communicated in sign language, responses were recorded in writing. For the analyses of student enumeration processes, the sets with one dot were not included since the enumeration processes can only be meaningfully differentiated from 2 points upwards.

Thus, 32 tasks per student were analyzed, resulting in a total of 7,264 tasks. To analyze student enumeration processes, we used heatmaps provided by the Tobii Pro Lab software together with AI (ML).
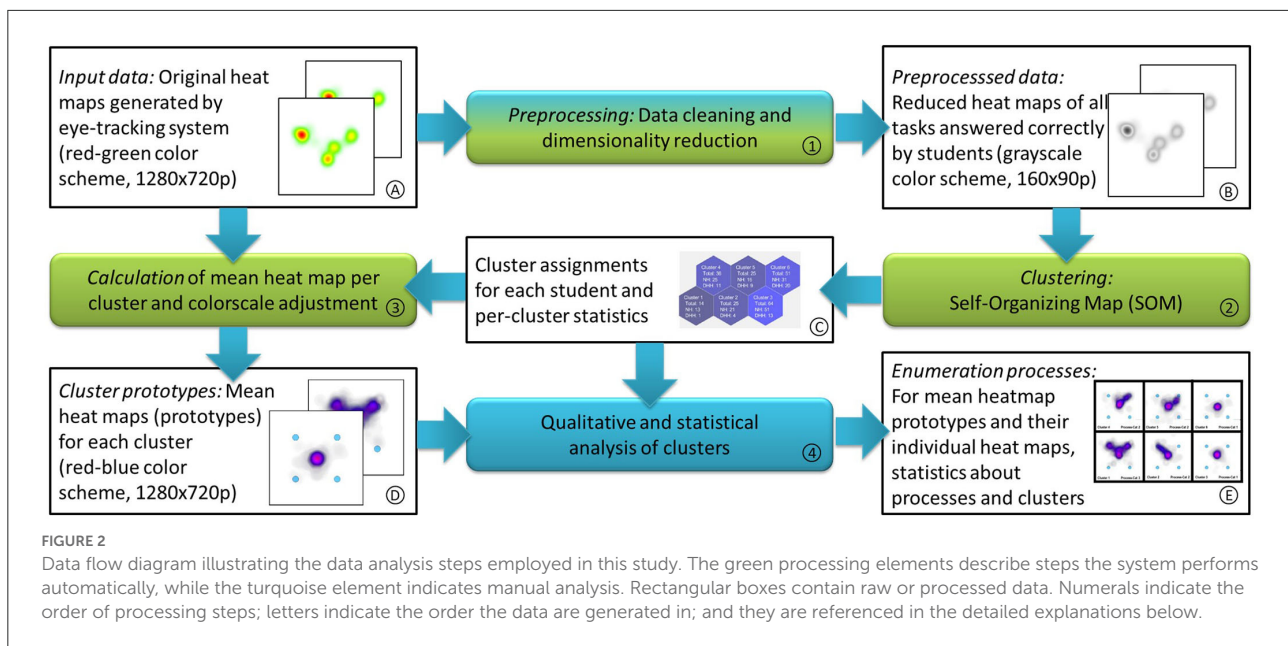
Data analysis

Input data

ET devices, like the one used in this study, offer different representations for the recorded ET data, such as detailed scan paths comprised of saccades and fixation times. A slightly less complex but rather intuitive representation is an aggregated heatmap, where the spatial distribution of gaze targets is represented as an image. This representation loses parts of the temporal information completely, such as the order in which the gaze shifts; and it represents other aspects only in relative terms, such as the time the gaze focuses on a single position. This information is color-coded relative to the length of the complete trial. Despite these shortcomings, patterns are easily accessible for human experts to interpret on heatmaps, and they allow the use of tried and tested image processing methods and ML approaches. Furthermore, previous studies (Schindler et al., 2019a,b) indicate the usefulness of heatmaps for assessing the students' enumeration processes since the enumeration process employed by the student can be identified by a human expert (Schindler et al., 2020a). In our processing pipeline, which is shown in Figure 2, we use the heatmaps generated by the Tobii Pro Lab Software as input data (Figure 2A). Each task of each student generates a single heatmap of 1,920x760 color pixels (~1.5 megapixels) for a total of 7,264 heatmaps with more than 10 billion data points (10,930,867,200).

Preprocessing

In an initial data cleaning step, heatmaps were removed from the data pool if the student did not solve the enumeration task correctly (Figure 2.1). This step is

FIGURE 2

Data flow diagram illustrating the data analysis steps employed in this study. The green processing elements describe steps the system performs automatically, while the turquoise element indicates manual analysis. Rectangular boxes contain raw or processed data. Numerals indicate the order of processing steps; letters indicate the order the data are generated in; and they are referenced in the detailed explanations below.

necessary to exclude a random gaze pattern where the student just guessed instead of enumerating the dots shown. After performing this filter step, 6,615 usable heatmaps remained.

Each heatmap was reduced in size by converting them to grayscale images and subsampling them to 160 x 90 pixels each (Figure 2B). This drastic reduction in dimensionality (by a factor of more than 300) without losing important information is possible since the heatmap images do not contain high-frequency components (that is, details on the level of a few pixels). Even isolated fixations are represented as smooth blobs, and the minimum blob size is large enough to always be conserved in the reduced-size images. This dimensionality reduction was done for computational reasons only. Executing the clustering algorithm on full-size images will yield very similar if not identical cluster assignments, but the time required on off-the-shelf computer hardware increases significantly: less than an hour to process all tasks with dimensionality-reduced images (160 x 90 gray value pixels) and an estimated week for original-sized images (1920 x 760 color pixels).

Clustering

The purpose of a clustering algorithm is to automatically assign multivariate data into subgroups (Everitt et al., 2011). In this study, each heatmap was grouped with similar heatmaps to automatically identify clusters, in which all heatmaps represent similar enumeration processes of the students (Figure 2.2). A cluster can be represented by its prototype, which, in this case, is the average heatmap (see Figure 3) of all the original heatmaps assigned to this cluster. Finding these groups automatically

required two decisions by the system designer: First, the concept of (dis-) similarity for the data had to be defined. This relates to the question, "when are two heat maps similar to each other [see (1) in the following]?" Second, the specific clustering algorithm had to be chosen, and an assumption about the number of clusters or granularity of the data had to be made [see (2) in the following].

(1) Defining the similarity between two images is broadly explored in the digital image processing community (Goshtasby, 2012). While there is often the possibility to have domain- or application-specific definitions, we opted for a standard approach in this study and used the Euclidean distance d between two heatmaps as a measure for their dissimilarity. It was calculated by comparing two heatmaps H_1 and H_2 pixel by pixel, that is, summing the squared pixel differences:

$$d = \|H_1 - H_2\|_2 = \sqrt{\sum_{i=1}^{160} \sum_{j=1}^{90} [H_1(i,j) - H_2(i,j)]^2}$$

Note that $H_1(i, j)$ describes the color of the pixel at position i, j in heatmap H_1. The distance d for any two identical heatmaps is always zero, and the more the pixel values differ, the larger d becomes. Likewise, the pattern recognition and machine learning fields have spawned large numbers of different clustering algorithms (Everitt et al., 2011), all with different advantages and disadvantages of their own. As we are interested in explorative data analysis, this study applies the SOM algorithm (Kohonen, 2001). The distinctive feature of SOMs in comparison to other cluster algorithms is the *a priori* assumption of a grid structure defining a relationship

between the clusters. This 2D-grid, from which the map part of the name is derived, can provide additional information for interpreting the clustering results. The algorithm contains an iterative optimization procedure, which randomly selects a heatmap and calculates the distance to all cluster prototypes. The closest prototype will then move toward the current heatmap, making the cluster prototype more similar to the heatmap. The moving cluster representatives' neighbors, as defined by the *a priori* grid, will also move in that direction, but to a lesser degree. Then, the next heatmap is chosen, and the procedure repeats until the optimization process converges, that is, the cluster centers stop moving when all similar heatmaps are assigned to the same cluster. Due to its nature, the algorithm may produce different results depending on the random initialization of cluster representatives, and the order the algorithm selects the individual heatmaps for comparison. For the algorithmic details, we refer to Kohonen (2001). For the implementation, we used the Matlab's Deep Learning Toolbox V. 13.0 (RRID: SCR_001622) with standard learning parameters.

(2) SOMs, like many other clustering algorithms, require choosing the number of groups and their relationships to each other as part of the algorithm. This can be done automatically by defining a cluster quality measure and performing an additional model selection optimization loop, which increases computational costs significantly, or it can be done manually by either guessing or choosing the number based on the domain knowledge. Since a previous study (Schindler et al., 2020b) identified three different processes, this study used six clusters in a 3 x 2 grid structure. Under the assumption that all the three different processes are present in the data and sufficiently dissimilar as measured by the Euclidian distance, they each form one group, and the additional groups can represent outliers, such as the previously unobserved processes or the subgroups of one of the three identified processes.

Cluster assignment

As a result of the clustering process, each heatmap is assigned to one of the six clusters (Figure 2C). The unused information in the clustering process can now be determined for each cluster. This includes the number of students who are DHH and hearing students and other statistics for each cluster.

Cluster prototypes

Since the clustering itself was performed on the size-reduced heatmaps, in the next step, full-size cluster prototypes (Figure 2.3) are calculated from the original heatmaps. To help with the visual inspection of the cluster prototypes and for easier distinction, we shift the color space for those prototypes from the red-green color table created by the ET software to a red-blue color table for the cluster prototypes. This can be done for RGB

images by copying the green channel into the blue channel and then setting the green channel to zero (Figure 2D).

Qualitative analysis

The unsupervised learning results in six clusters each for every task (item) used in this study. For each of the non-empty clusters, we interpreted the average heatmaps (Figure 2.4). We assigned categories of processes based on a system that had been developed inductively in an earlier study (Schindler et al., 2020b). For example, for the dice patterns, we distinguished the following three *types of processes*:

(1) Simultaneous enumeration: Presented dots are enumerated simultaneously. The gazes are predominantly in the middle of the pattern (on the middle point).
(2) Enumeration through the use of groups/structures: Presented dots are enumerated through the use of groups/ structures: Gazes are on parts of the dots, for example, to one side, indicating the use of symmetries.
(3) Enumeration through counting: At least half of the dots are looked at—with gazes on at least half of the points—or every dot is counted, which means that gazes are on every dot.

Note that for the random arrangements in the counting range, the three types of processes were minimally different: *(1) Quasi-simultaneous enumeration*, *(2) Enumeration through the use of groups/structures*, and *(3) Enumeration through counting*. The first process in this case was the quasi-simultaneous enumeration because sets of dots in the counting range cannot be enumerated simultaneously but quasi-simultaneously (see Schindler et al., 2020b for more detail).

In practice, this means that, for the example of the dice five, we assigned type 1 to C3 and C6, type 2 to C2, C4, and C5, and type 3 to C1 (Figure 3).

Two of our researchers categorized all average heatmaps in this way. We assigned types of processes for each of the four conditions: (a) subitizing range (2–4) in random arrangements, (b) counting range (5–9) in random arrangements, (c) dice range (2–6) in canonical arrangements, and (d) beyond dice range (7–9) in canonical arrangements.

We then calculated how many student heatmaps were assigned to each type of process. For this, we drew on the information of how many heatmaps (of students who are DHH and of hearing students) were in each cluster. For example, for the dice five, we found (1) *type 1*: 33 (=13+20) DHH and 82 (=51+31) hearing, (2) *type 2*: 24 DHH and 62 hearing, and (3) *type 3*: 1 DHH and 13 hearing.

Statistical analysis

To identify group differences in student enumeration processes, we carried out k x 2 chi square tests, which allowed us to compare two independent samples for a k-stepped property

FIGURE 3
Example of clusters for the dice five (canonical arrangement, dice range).
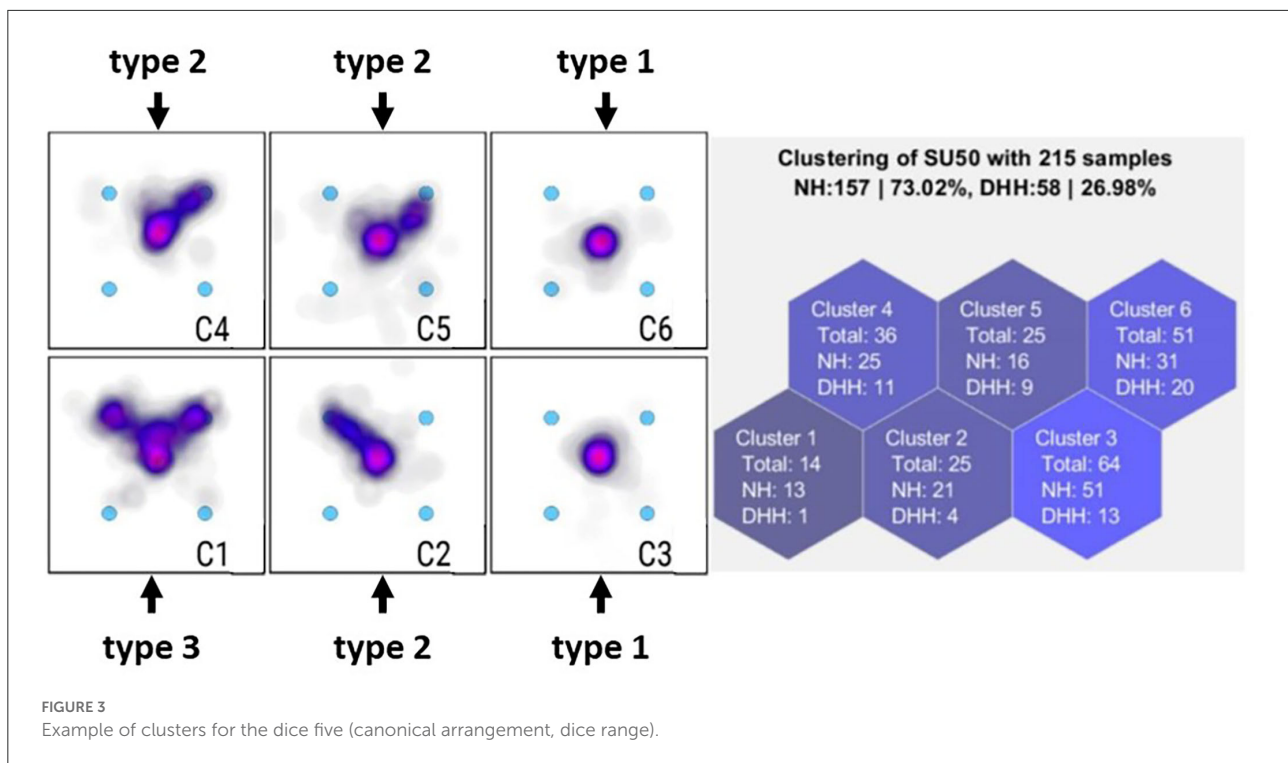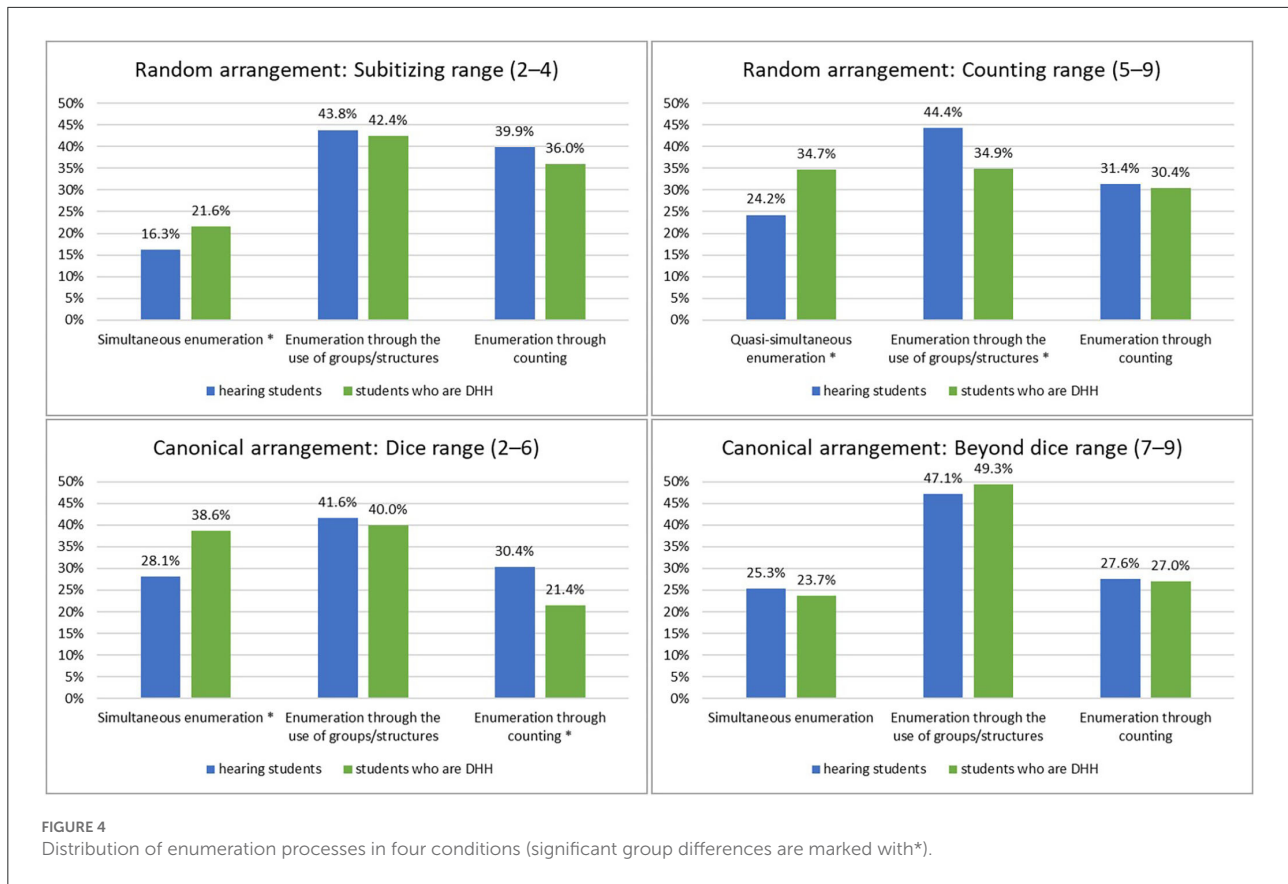
(Figure 2.4). In our case, we had two groups (DHH, hearing) and three enumeration processes. We conducted a 3 x 2 chi square test for each condition: (a) subitizing range (2–4) in random arrangements, (b) counting range (5–9) in random arrangements, (c) dice range (2–6) in canonical arrangements, and (d) beyond dice range (7–9) in canonical arrangements. We calculated effect sizes using *Cramérs V*. According to Cohen (1988), *Cramérs V* can be interpreted as follows: $V = 0.10$ is a small effect, $V = 0.30$ is a medium effect, and $V \geq 0.50$ is a large effect. For degrees of freedom of 2, like in our study, according to Cohen (1988) $V = 0.07$ is a small effect, $V = 0.21$ a medium effect, and ≥ 0.35 a large effect. Within the subitizing range, 1,961 items were analyzed (2–4 in three different arrangements); within the counting range, 2,972 items were analyzed (5–9 in three different arrangements); in the dice range, 1,084 items were analyzed (2–6 in one arrangement); and in the beyond dice range, 598 items were analyzed (7–9 in one arrangement). In the cases where the chi square test showed significant group differences, we calculated cell tests to investigate what the significant group differences were due to. We carried out the statistical analyses using the statistics and analysis software IBM SPSS 28 (RRID: SCR_019096).

Results

Chi square tests were performed to investigate the relationship between hearing loss and enumeration processes.

The chi square tests revealed significant differences[1] in the distribution of enumeration processes (Figure 4) between students who are DHH and hearing students for both conditions in random arrangements: (a) subitizing range: χ^2 (2) = 7.91, $p = 0.038$, $V = 0.06$ and (b) counting range: χ^2 (2) = 35.54, $p < 0.001$, $V = 0.11$. Effect sizes were small. For the dice range in the canonical arrangement condition (c), chi square test also revealed significant differences with small effect sizes in the distribution of enumeration processes between DHH and hearing students: χ^2 (2) = 13.87, $p = 0.002$, $V = 0.11$. Yet, no significant difference was found for the larger sets in canonical arrangement (beyond dice range): χ^2 (2) = 0.26, $p = 0.878$. This means that the distribution of enumeration processes did not differ between students who are DHH and hearing students in the beyond dice range in the canonical arrangement condition. Cell tests for the differences between groups revealed the following results: *(a) Subitizing range:* Students who are DHH used simultaneous enumeration more often than hearing students (χ^2 (1) = 7.53, $p = 0.018$, $V = 0.06$). *(b) Counting range:* Students who are DHH used quasi-simultaneous enumeration more often than hearing students (χ^2 (1) = 31.74, $p < 0.001$, $V = 0.10$). Students who are DHH used enumeration of groups less often than hearing students (χ^2 (1) = 20.95, $p < 0.001$, $V = 0.08$). *(c) Dice range in canonical arrangement:* Students who are DHH used

---

1 Bonferroni-Holm adjusted *p*-values are given in the following.

FIGURE 4
Distribution of enumeration processes in four conditions (significant group differences are marked with*).

simultaneous enumeration more often than hearing students (χ^2 (1) = 11.04, $p < 0.001$, $V = 0.10$). Students who are DHH used enumeration through counting less often compared to hearing students (χ^2 (1) = 8.49, $p = 0.008$, $V = 0.09$). Also here, all effect sizes were small.

Summary and discussion

The aim of this study is to investigate if students who are DHH differ from hearing students in small number enumeration processes. We used the ET data of 227 students (164 hearing students, 63 students who were DHH). For analyzing the students' eye movements, we used ET in combination with AI, specifically a clustering algorithm, and the analysis of human experts to identify enumeration processes from the heatmaps of student gaze distributions. Our study investigates the enumeration processes of school-age students who are DHH in a large-scale study—and it brought interesting results to light, indicating that students who are DHH used more advantageous processes than their hearing peers in different task conditions. A further feature of this work is that we used AI to analyze the differences in small number enumeration processes between students who are DHH and hearing students. The ML-based processing pipeline introduced in this study, first and foremost,

allowed us to efficiently process gaze heat maps for a large number of students by automatically clustering the enumeration strategies based on the patterns in the data. The automated clustering analysis also provided an independent view of the data that can support human researchers as an "AI colleague" in the interpretation of gaze data. The presented algorithms are not specific to the task of small set enumeration but can be applied to any set of heat maps. Hence, the pipeline can serve as a blueprint for similar studies, but also as a basis for the development of automated student assessment tools supporting researchers and teachers.

We examined student enumeration processes of sets under four conditions: both the subitizing range (2–4) and the counting range (5–9) in random arrangements, and both the dice range (2–6) and beyond dice range (7–9) in canonical arrangements. We found significant differences in the distribution of enumeration processes between students who are DHH and hearing students in both random arrangement conditions: In the subitizing range and the counting range, students who are DHH used more advantageous enumeration processes than hearing students. This means that students who are DHH enumerated the presented dots more often (quasi-) simultaneously, that is, they took in all dots at a glance. For dice patterns in the canonical arrangement, we also found significant differences in the distribution of enumeration

processes between students who are DHH and hearing students. Also here, students who are DHH more often enumerated the presented items simultaneously, indicating that students who are DHH recognized and recalled the dice patterns more often than hearing students. However, it should be noted that the effect sizes were small for all significant differences; practically, the effect of group on students' enumeration processes is small. Yet, the significant differences are interesting against the backdrop of the following aspect: The sample of students who are DHH in our study was diverse, for example, regarding language use, which is known to be a major influencing factor on the learning and achievement of students who are DHH. Also, approximately two-thirds of the students who are DHH (65%) in our study had mathematical difficulties. With a large heterogeneity within this group in terms of language use and hearing loss and the fact that a substantial number of them had MD, it is remarkable that, still, significant group differences were found—with students who are DHH performing better than hearing students. Even if effect sizes are small, our findings indicate that students who are DHH in third to fifth grade do not appear to have difficulties in small number enumeration processes as compared to hearing students at that age but, rather, appear to use more advantageous enumeration processes than their hearing peers. They use subitizing and groupitizing more often, and they recognize and recall dice patterns more often. This is interesting and contributes to the state of research in different ways.

Bull et al. (2006) had shown that, for adults who are DHH and at a high academic performance level (university students), there are no differences in response times in small number enumeration as compared to their hearing peers. Our study showed that children who are DHH at the age of approximately 10–11 years not only perform as well as their hearing peers, but they partially appear to perform even better than their hearing peers in that specific area. This is interesting since also in our sample, the students who are DHH had significant difficulties in mathematics. These difficulties do not seem to go hand-in-hand with subitizing deficits—unlike students with MD, where a "dysfunctional subitizing mechanism" (Schleifer and Landerl, 2011, p. 280) appears to accompany MD. Our results indicate that the difficulty profile of students who are DHH appears to differ from that of students with MD in general.

Furthermore, our study found that the students who were DHH also performed better than their hearing peers in conceptual subitizing, also called groupitizing (Clements, 1999; Starkey and McCandliss, 2014): In the counting range, students who are DHH used the quasi-simultaneous enumeration more often, where it is necessary to use the part-whole relationship and where conceptual understanding is necessary (Krajewski and Schneider, 2009; Starkey and McCandliss, 2014). This is particularly interesting since students who are DHH generally perform lower than their hearing peers in mathematics. Still, our study indicates that they do have this fundamental conceptual knowledge and apply it even more often than their hearing peers.

Finally, it is worth mentioning that students who are DHH recognized and recalled dice patterns at one glance more often than their hearing peers. Less often, they needed to take in the dots serially. We can only speculate about the reasons. One possibility is that students who are DHH are more likely to memorize and recall visual patterns. Overall, it is often discussed that visual abilities are enhanced in relation to auditory deprivation in people who are DHH (Hauthal et al., 2013). Many studies have implied that a lack of auditory experience from an early age influences the organization of the human brain for peripheral and central visual processing and is often compensated with supranormal performance in other sensory systems, such as vision (Alencar et al., 2019). Scott et al. (2014) found that adults with profound, congenital, and hereditary hearing loss have a network of brain regions exhibiting enhanced responsiveness to peripheral visual stimuli. Furthermore, deaf persons appear to have a different visual viewing behavior in some aspects, for example, a preferential central fixation pattern compared to hearing persons (Lao et al., 2017). However, this link to a better visual ability and performance has predominantly been associated with sign language, as some advantages in visual-spatial tasks are not only found in deaf individuals but also in hearing individuals who are skilled signers (Marschark et al., 2015). However, in our study group, most of the students who were DHH used spoken language or sign-supported speech as a preferred communication method, not sign language. Yet, in their regular schooling, they were in contact with sign language frequently, and it is further possible that visual orientation played an important adaptive or compensational role for these students, as well as in their teaching. Didactics and teaching materials in the special schools for students who are DHH are often supported by visualizations in such a way that they can be easily accessed by all students, both sign language and spoken language users (Nunes, 2004). Knoors and Marschark (2014) emphasized in their research that students who are DHH are not hearing learners who simply cannot hear; they may utilize different abilities in dealing with tasks.

Only in the canonical arrangements beyond the dice range did the students who are DHH and the hearing students apply processes similarly. One possible explanation is that here, the hearing students had more experience dealing with symmetries, and this compensated for the advantages that students who are DHH may have had otherwise.

Limitations and implications for future studies

Besides summarizing and discussing our results, we want to discuss the *limitations* of our study and point out some suggestions for *future work* that emerge. One possible limitation of this study lies in the selection and composition of the

sample. The group of hearing students consisted of fifth-grade students (mean: 10.9 years, SD: 0.7 years), and the group of students who are DHH consisted of fifth-, fourth-, and third-grade students (mean: 11.9 years, SD: 1.0). So, students who are DHH in our study were on average in lower grade levels, but at the same time older than the hearing students, which makes it difficult to compare the groups in relation to their mathematical development in general. However, given that children learn to enumerate small numbers of dots well before the age of 10—before or at the very beginning of school—and since the curriculum in grades 3–5 does not cover small number enumeration, the grade levels were not critical for this skill. Therefore, the heterogeneity of the groups was acceptable and—in our view—biased the results marginally at most. Furthermore, the group of students who are DHH showed a large heterogeneity in terms of familial first language (L1), their communication modality, the degree of hearing loss, and their use of hearing aids or cochlear implants. On the other hand, the composition of the group of students who are DHH in our study represented well the heterogeneity of students who are DHH that is often encountered in special schools for students who are DHH in Germany: The DHH sample in our study had the commonality that they attended special schools for students who are DHH in Germany and they all had a peripheral hearing loss and, in a few cases, a central auditory processing disorder, yet, they were diverse otherwise. Subgroups were small, which prevented us from running subgroup tests. However, for future research, it would be valuable to look closer into language abilities and group differences within the group of students who are DHH in terms of communication modality, degrees of hearing loss, etc.

Regarding the applied clustering methodology, it is worth noting that the generic approach we used in this article was not optimized for this application. While we do not expect substantially different results if a different clustering algorithm had been applied, the choice of the right dissimilarity measure might be more sensitive in the future. In this study, we used the Euclidean distance, which may not be optimal in cases where the appearance of the heatmaps that belong to one enumeration process is, in part, the same as the appearance of the heatmaps belonging to another process. In such cases, the heatmaps that correspond to two different processes will be scored to be similar and might end up in the same cluster. Hence, investigating distance metric learning methods for eye-tracking heatmaps that provide a stronger separation for those cases is a promising topic to investigate in the future (see Xing et al., 2002).

Conclusion

The results of this study indicate that students who are DHH do not lag in the enumeration of small numbers compared to hearing students but rather appear to use more advantageous processes in enumerating small sets than their hearing peers. Our study suggests that difficulties in enumerating small numbers do not coincide with the MD that students who are DHH often have, which offers interesting perspectives for further research.

Pedagogical implications

For educational practice, our results indicate that students who are DHH—regardless of their language modality—appear not to need support in enumeration processes, such as the part-whole relationship. In this respect, the need for support differs from that necessary for students with MD in general, where the practice of, for example, basic visual small number enumeration processes can be valuable, especially in early grades. In our study, even students who were DHH and had MD showed advanced skills in small number enumeration. For supporting students who are DHH, a focus on other aspects of mathematics learning that are known to be difficult for DHH learners, including conceptual understanding, problem-solving, and word problem solving, appear to be more significant. These results and consequences are to be taken into consideration for teacher training in special education, ideally across all spoken and signed languages, which should be guided by the needs and strengths of the group of students who are DHH. Some pedagogical suggestions for DHH teachers and early intervention practitioners "from research to practice" have been recently made; for example, on teaching numeracy and early math concepts (Kritzer and Green, 2021), fostering fraction learning (Mousley, 2021), spatial reasoning (Thom and Hallenbeck, 2021), and usage of sign language in the mathematics classroom (Krause and Wille, 2021). Our study indicates that small number enumeration appears to be among the relative strengths of students who are DHH. This is encouraging given the general difficulties and delays in their mathematical development.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

MS and KS contributed to the conception and design of the study. MS, JD, and AS conducted the study and collected the data. MS, JD, AS, ES, and AL performed the data analysis. MS, AS, ES, AL, and KS wrote the manuscript. All authors read and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Aldenderfer, M. S., and Blashfield, R. K. (1984). *Cluster Analysis.* Newbury Park,CA: Sage. doi: 10.4135/9781412983648

Alencar, C. D. C., Butler, B. E., and Lomber, S. G. (2019). What and how the deaf brain sees. *J. Cogn. Neurosci.* 31, 1091–1109. doi: 10.1162/jocn_a_01425

Anobile, G., Castaldi, E., Maldonado Moscoso, P. A., Burr, D. C., and Arrighi, R. (2020). Groupitizing: a strategy for numerosity estimation. *Sci. Rep.* 10:13436. doi: 10.1038/s41598-020-68111-1

Ansell, E., and Pagliaro, C. M. (2006). The relative difficulty of signed arithmetic story problems for primary level deaf and hard-of-hearing students. *J. Deaf Stud. Deaf Educ.* 11, 153–170. doi: 10.1093/deafed/enj030

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

Austin, G. F. (1975). Knowledge of selected concepts obtained by an adolescent deaf population. *Am. Ann. Deaf* 0.120, 360–370.

Barbosa, H. H. (2014). Early mathematical concepts and language: a comparative study between deaf and hearing children. *Educ. Pesqui.* 40, 163–179. doi: 10.1590/S1517-97022014000100011

Bedoya-Ríos, N. M., and Dorneles, B. V. (2021). "Strategies and accuracy in the number line task in Colombian and Brazilian deaf children," in *Mathematical Reasoning of Children and Adults – Teaching and Learning From an Interdisciplinary Perspective*, eds. A. G. Spinillo, S. L. Lautert, and R. E. de Souza Rosa Borba (Cham: Springer), 95–111. doi: 10.1007/978-3-030-69657-3_5

Blatto-Vallee, G., Kelly, R. R., Gaustad, M. G., Porter, J., and Fonzi, J. (2007). Visual-spatial representation in mathematical problem solving by deaf and hearing students. *J. Deaf Stud. Deaf Educ.* 12, 432–448. doi: 10.1093/deafed/enm022

Bull, R. (2008). "Deafness, numerical cognition, and mathematics," in *Deaf Cognition: Foundations and Outcomes*, eds. M. Marschark, and P. Hauser (Oxford: Oxford University Press), 170–200. doi: 10.1093/acprof:oso/9780195368673.003.0006

Bull, R., Blatto-Vallee, G., and Fabich, M. (2006). Subitizing, magnitude representation, and magnitude retrieval in deaf and hearing adults. *J. Deaf Stud. Deaf Educ.* 11, 289–302. doi: 10.1093/deafed/enj038

Bull, R., Marschark, M., Sapere, P., Davidson, W. A., Murphy, D., and Nordmann, E. (2011). Numerical estimation in deaf and hearing adults. *Learn. Individ. Differ.* 21, 453–457. doi: 10.1016/j.lindif.2011.02.001

Butterworth, B. (2005). "Developmental dyscalculia," in *Handbook of Mathematical Cognition*, ed. J. I. D. Campbell (New York: Psychology Press), 455–467.

Carrasco, M. (2011). Visual attention: the past 25 years. *Vis. Res.* 51, 1484–1525. doi: 10.1016/j.visres.2011.04.012

Clarke, B., Cheeseman, J., and Clarke, D. (2006). The mathematical knowledge and understanding young children bring to school. *Math. Educ. Res. J.* 18, 78–102. doi: 10.1007/BF03217430

Clements, D. H. (1999). Subitizing: what is it? why teach it? *Teach. Child. Math.* 5, 400–405. doi: 10.5951/TCM.5.7.0400

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences, 2nd Edn.* Hillsdale, N.J.: L. Erlbaum Associates.

Department of Education (2013). *Mathematics programmes of study: Key stages 1 and 2 National curriculum in England.* Available online at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/335158/PRIMARY_national_curriculum_-_Mathematics_220714.pdf (accessed February 15, 2022).

Edwards, A., Edwards, L., and Langdon, D. (2013). The mathematical abilities of children with cochlear implants. *Child Neuropsychol.* 19, 127–142. doi: 10.1080/09297049.2011.639958

Everitt, B., Landau, S., Leese, M., and Stahl, D. (2011). *Cluster Analysis*, 5th Edn. Chichester: Wiley. doi: 10.1002/9780470977811

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

Frostad, P. (1996). Mathematical achievement of hearing impaired students in Norway. *Eur. J. Spec. Needs Educ.* 11, 67–81. doi: 10.1080/0885625960110105

Garrote, A., Moser Opitz, E., and Ratz, C. (2015). Mathematische Kompetenzen von Schülerinnen und Schülern mit dem Förderschwerpunkt geistige Entwicklung. *Eine Querschnittstudie. Emp. Sonderpäd.* 7, 24–40.

Geary, D. C. (2013). Early foundations for mathematics learning and their relations to learning disabilities. *Curr. Dir. Psychol. Sci.* 22, 23–27. doi: 10.1177/0963721412469398

Gelman, R., and Gallistel, C. R. (1986). *The Child's Understanding of Number.* Cambridge, MA: Harvard University Press. doi: 10.4159/9780674037533

Goshtasby, A. A. (2012). *Image Registration.* London: Springer. doi: 10.1007/978-1-4471-2458-0

Gottardis, L., Nunes, T., and Lunt, I. (2011). A synthesis of research on deaf and hearing children's mathematical achievement. *Deafness Educ. Int.* 13, 131–150. doi: 10.1179/1557069X11Y.0000000006

Gray, S. A., and Reeve, R. A. (2014). Preschoolers' dot enumeration abilities are markers of their arithmetic competence. *PLoS ONE* 9:e99428. doi: 10.1371/journal.pone.0094428

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

Hannula-Sormunen, M. M., Lehtinen, E., and Räsänen, P. (2015). Preschool children's spontaneous focusing on numerosity, subitizing, and counting skills as predictors of their mathematical performance seven years later at school. *Math. Think. Learn.* 17, 155–177. doi: 10.1080/10986065.2015.1016814

Hauthal, N., Sandmann, P., Debener, S., and Thorne, J. D. (2013). Visual movement perception in deaf and hearing individuals. *Adv. Cogn. Psychol.* 9, 53–61. doi: 10.5709/acp-0131-z

Hecht, T., Sinner, D., Kuhl, J., and Ennemoser, M. (2011). Differenzielle effekte eines trainings der mathematischen basiskompetenzen bei kognitiv schwachen

grundschülern und schülern der förderschule mit dem schwerpunkt lernen-reanalyse zweier studien. *Emp. Sonderpäd.* 3, 308–323.

Heiling, K. (1995). *The development of deaf school children.* Academic achievement and social processes. Signum.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends Cogn. Sci.* 7, 498–504. doi: 10.1016/j.tics.2003.09.006

Henner, J., Pagliaro, C., Sullivan, S., and Hoffmeister, R. (2021). Counting differently: assessing mathematics achievement of signing deaf and hard of hearing children through a unique lens. *Am. Ann. Deaf* 166, 318–341. doi: 10.1353/aad.2021.0023

Hine, W. D. (1970). The attainment of children with partial hearing. *J. Br. Assoc. Teach. Deaf* 68, 129–135.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., and Van de Weijer, J. (2011). *Eye Tracking: A Comprehensive Guide to Methods and Measures.* Oxford: Oxford University Press.

Hsin, C. -Y., Lo, Y. -H., and Tseng, P. (2021). Effect of non-canonical spatial symmetry on subitizing. *Front. Psychol.* 12:562762. doi: 10.3389/fpsyg.2021.562762

Hyde, M., Zevenbergen, R., and Power, D. (2003). Deaf and hard of hearing students' performance on arithmetic word problems. *Am. Ann. Deaf* 148, 56–64. doi: 10.1353/aad.2003.0003

Just, M. A., and Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychol. Rev.* 87, 329–354. doi: 10.1037/0033-295X.87.4.329

Kelly, R., and Mousley, K. (2001). Solving word problems: more than reading issues for deaf students. *Am. Ann. Deaf* 146, 251–262. doi: 10.1353/aad.2012.0088

Knoors, H., and Marschark, M. (2014). *Teaching Deaf Learners: Psychological and Developmental Foundations.* Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199792023.001.0001

Kohonen, T. (2001). *Self-Organizing Maps*, 3rd Edn. Berlin: Springer. doi: 10.1007/978-3-642-56927-2

Krajewski, K., and 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. *Learn. Instr.* 19, 513–526. doi: 10.1016/j.learninstruc.2008.10.002

Krause, C. M., and Wille, A. M. (2021). Sign language in light of mathematics education: an exploration within semiotic and embodiment theories of learning mathematics. *Am. Ann. Deaf* 166, 352–377. doi: 10.1353/aad.2021.0025

Kreilinger, I. L., Roesch, S., Moeller, K., and Pixner, S. (2021). Mastery of structured quantities like finger or dice patterns predict arithmetic performance. *Cogn. Process.* 22, 93–104. doi: 10.1007/s10339-020-00994-4

Kritzer, K. L. (2008). Family mediation of mathematically based concepts while engaged in a problem-solving activity with their young deaf children. *J. Deaf Stud. Deaf Educ.* 13, 503–517. doi: 10.1093/deafed/enn007

Kritzer, K. L. (2009). Barely started and already left behind: a descriptive analysis of the mathematics ability demonstrated by young deaf children. *J. Deaf Stud. Deaf Educ.* 14, 409–421. doi: 10.1093/deafed/enp015

Kritzer, K. L., and Green, L. (2021). Playing with code: an innovative way to teach numeracy and early math concepts to young DHH children in the 21st century. *Am. Ann. Deaf* 166, 409–423. doi: 10.1353/aad.2021.0027

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

Lao, J., Stoll, C., Dye, M., Pascalis, O., and Caldara, R. (2017). Deafness amplifies visual information sampling during face recognition. *J. Vis.* 17:24. doi: 10.1167/17.10.24

Leigh, G. and Marschark, M., (2016). "Recognizing diversity in deaf education: from Paris to Athens with a diversion to Milan," in *Diversity in Deaf Education,* eds. M. Marschark, V. Lampropoulou, and E. K. Skordilis (New York: Oxford University Press), 1–20. doi: 10.1093/acprof:oso/9780190493073.003.0001

Leybaert, J., and Van Custem, M. -N. (2002). Counting in sign language. *J. Exp. Child Psychol.* 81, 482–501. doi: 10.1006/jecp.2002.2660

Lilienthal, A. J., and Schindler, M. (2019). "Eye tracking research in mathematics education: A PME literature review," in *Proceedings of the 43rd Conference of the International Group for the Psychology of Mathematics Education*, Vol. 4, eds. M. Graven, H. Venkat, A. Essien, and P. Vale (Pretoria: PME), 62.

Mandler, G., and Shebo, B. J. (1982). Subitizing: an analysis of its component processes. *J. Exp. Psychol. Gen.* 111, 1–22. doi: 10.1037/0096-3445.111.1.1

Marschark, M., and Knoors, H. (2012). Educating deaf children: language, cognition, and learning. *Deafness Educ. Int.* 14, 136–160. doi: 10.1179/1557069X12Y.0000000010

Marschark, M., and Leigh, G. (2016). "Recognizing diversity in deaf education: now what do we do with it?!" in *Diversity in Deaf Education*, eds. M. Marschark, V. Lampropoulou, and E. K. Skordilis (New York, NY: Oxford University Press), 507–536. doi: 10.1093/acprof:oso/9780190493073.003.0018

Marschark, M., Morrison, C., Lukomski, J., Borgna, G., and Convertino, C. (2013). Are deaf students visual learners? *Learn. Individ. Differ.* 1, 156–162. doi: 10.1016/j.lindif.2013.02.006

Marschark, M., Spencer, L. J., Durkin, A., Borgna, G., Convertino, C., Machmer, E., et al. (2015). Understanding language, hearing status, and visual spatial skills. *J. Deaf Stud. Deaf Educ.* 20, 310–330. doi: 10.1093/deafed/env025

Mitchell, T. M. (1997). *Machine Learning*. New York, NY: McGraw Hill.

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

Moeller, K., Neuburger, S., Kaufmann, L., Landerl, K., and Nuerk, H. C. (2009). Basic number processing deficits in developmental dyscalculia: evidence from eye tracking. *Cogn. Dev.* 24, 371–386. doi: 10.1016/j.cogdev.2009.09.007

Moser Opitz, E., Freesemann, O., Prediger, S., Grob, U., Matull, I., and Hußmann, S. (2017). Remediation for students with mathematics difficulties: an intervention study in middle schools. *J. Learn. Disabil.* 50, 724–736. doi: 10.1177/0022219416668323

Mousley, K. (2021). Pedagogical suggestions to foster fraction learning. *Am. Ann. Deaf* 166, 342–351. doi: 10.1353/aad.2021.0024

Mousley, K., and Kelly, R. R. (2018). Developing deaf students fraction skills requires understanding magnitude and whole number division. *J. Educ. Learn.* 7, 12–20. doi: 10.5539/jel.v7n2p12

Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective.* Cambridge, MA: The MIT Press.

Nunes, T. (2004). *Teaching Mathematics to Deaf Children.* London: Whurr.

Nunes, T., and Moreno, C. (2002). An intervention program for promoting deaf pupils' achievement in mathematics. *J. Deaf Stud. Deaf Educ.* 7, 120–133. doi: 10.1093/deafed/7.2.120

Pagliaro, C. M. (2015). "Developing numeracy in individuals who are deaf and hard of hearing," in *Educating Deaf Learners: Creating a Global Evidence Perspective*, eds. H. Knorrs, and M. Marschark (Oxford: Oxford University Press), 173–195. doi: 10.1093/acprof:oso/9780190215194.003.0008

Pagliaro, C. M., and Kritzer, K. L. (2013). The math gap: a description of the mathematics performance of preschool-aged deaf/ hard-of-hearing children. *J. Deaf Stud. Deaf Educ.* 18, 139–160. doi: 10.1093/deafed/ens070

Pitters, J. (2018). *Faktoren der Entwicklung mathematischer Basiskompetenzen: Eine Untersuchung mit Schülerinnen und Schülern der fünften Jahrgangsstufe an Integrierten Gesamtschulen und Oberschulen [Factors in the development of basic mathematical skills: An investigation with fifth-grade students in integrated comprehensive schools and secondary schools].* Carl von Ossietzky Universität Oldenburg.

Posner, M. I. (1980). Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25. doi: 10.1080/00335558008248231

Qi, S., and Mitchell, R. E. (2012). Large-scale academic achievement testing of deaf and hard-of-hearing students: past, present, and future. *J. Deaf Stud. Deaf Educ.* 17, 1–18. doi: 10.1093/deafed/enr028

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychol. Bull.* 124, 372–422. doi: 10.1037/0033-2909.124.3.372

Santos, S., and Cordes, S. (2022). Math abilities in deaf and hard of hearing children: the role of language in developing number concepts. *Psychol. Rev.* 129, 199–211. doi: 10.1037/rev0000303

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

Schindler, M. (2021). "Eye-Tracking in der mathematikdidaktischen Forschung: Chancen und Herausforderungen," [Eye tracking in mathematics education research: Opportunities and challenges] in *Beiträge zum Mathematikunterricht 2021*, eds. K. Hein, C. Heil, S. Ruwisch, and S. Prediger (Münster: WTM), 21–28.

Schindler, M., Bader, E., Lilienthal, A. J., Schindler, F., and Schabmann, A. (2019a). Quantity recognition in structured whole number representations of students with mathematical difficulties: An eye-tracking study. *Learn. Disabil. A Contemp. J.* 17, 5–28.

Schindler, M., and Lilienthal, A. J. (2018). "Eye-tracking for studying mathematical difficulties-also in inclusive settings," in *Proceedings of the 42nd Conference of the International Group for the Psychology of Mathematics Education*,

Vol. 4, eds. E. Bergqvist, M. Österholm, C. Granberg, and L. Sumpter (Umeå: PME), 115–122.

Schindler, M., and Lilienthal, A. J. (2019). Domain-specific interpretation of eye tracking data: towards a refined use of the eye-mind hypothesis for the field of geometry. *Educ. Stud. Math.* 101, 123–139. doi: 10.1007/s10649-019-9878-z

Schindler, M., Schaffernicht, E., and Lilienthal, A. J. (2019b). "Differences in quantity recognition of students with and without mathematical difficulties: Analysis through eye-tracking and AI," in *Proceedings of the 43rd Conference of the International Group for the Psychology of Mathematics Education*, Vol. 3, eds. M. Graven, H. Venkat, A. Essien, and P. Vale (Pretoria: PME), 281–288.

Schindler, M., Schaffernicht, E., and Lilienthal, A. J. (2020a). "Identifying strategies through eye tracking and unsupervised learning: The case of quantity recognition," in *Proceedings of the 44th Conference of the International Group for the Psychology of Mathematics Education*, eds. M. Inprasitha, N. Changsri, and N. Boonsena (Khon Kaen: PME), 518–527.

Schindler, M., Schovenberg, V., and Schabmann, A. (2020b). Enumeration processes of children with mathematical difficulties: an explorative eye-tracking study on subitizing, groupitizing, counting, and pattern recognition. *Learn. Disabil. A Contemp. J.* 18, 193–211.

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

Scott, G. D., Karns, C. M., Dow, M. W., Stevens, C., and Neville, H. J. (2014). Enhanced peripheral visual processing in congenitally deaf humans is supported by multiple brain regions, including primary auditory cortex. *Front. Hum. Neurosci.* 26, 177. doi: 10.3389/fnhum.2014.00177

Spencer, P. E., and Marschark, M. (2010). *Evidence-Based Practice in Educating Deaf and Hard-Of-Hearing Students*. Oxford: Oxford University Press.

Starkey, G. S., and McCandliss, B. D. (2014). The emergence of "groupitizing" in children's numerical cognition. *J. Exp. Child Psychol.* 126, 120–137. doi: 10.1016/j.jecp.2014.03.006

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

Starkey, P., and Cooper, R. G. (1995). The development of subitizing in young children. *Br. J. Dev. Psychol.* 13, 399–420. doi: 10.1111/j.2044-835X.1995.tb00688.x

Strohmaier, A. R., MacKay, K. J., Obersteiner, A., and Reiss, K. (2020). Eye-tracking methodology in mathematics education research: a systematic literature review. *Educ. Stud. Math.* 104, 147–200. doi: 10.1007/s10649-020-09948-1

Swanwick, R., Oddy, A., and Roper, T. (2005). Mathematics and deaf children: an exploration of barriers to success. *Deafness Educ. Int.* 7, 1–21. doi: 10.1179/146431505790560446

Thom, J. S., and Hallenbeck, T. (2021). Beyond words/signs: DHH learners' spatial reasoning in mathematics as embodied cognition. *Am. Ann. Deaf* 166, 378–408. doi: 10.1353/aad.2021.0026

Titus, J. C. (1995). The concept of fractional number among deaf and hard of hearing students. *Am. Ann. Deaf* 140, 255–263. doi: 10.1353/aad.2012.0582

Tobii, A. B. (2019). *User's manual Tobii Pro X3-120 Eye Tracker.* Available online at: https://www.tobiipro.com/siteassets/tobii-pro/user-manuals/tobii-pro-x3-120-user-manual.pdf/?v=1.0.9 (accessed February 10, 2022).

Todman, J., and Seedhouse, E. (1994). Visual-action code processing by deaf and hearing children. *Lang. Cogn. Proc.* 9, 129–141. doi: 10.1080/01690969408402113

Traxler, C. (2000). The stanford achievement test, 9th edition: national norming and performance standards for deaf and hard-of-hearing students. *J. Deaf Stud. Deaf Educ.* 5, 337–348. doi: 10.1093/deafed/5.4.337

Van der Sluis, S., de Jong, P. F., and van der Leij, A. (2004). Inhibition and shifting in children with learning deficits in arithmetic and reading. *J. Exp. Child Psychol.* 87, 239–266. doi: 10.1016/j.jecp.2003.12.002

Viesel-Nordmeyer, N., Schurig, M., Bos, W., and Ritterfeld, U. (2019). Effects of pre-school mathematical disparities on the development of mathematical and verbal skills in primary school children. *Learn. Dis.: A Cont. J.* 17, 149–164.

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

Wege, T. E., Trezise, K., and Inglis, M. (2022). Finding the subitizing in groupitizing: Evidence for parallel subitizing of dots and groups in grouped arrays. *Psychon. Bull. Rev.* 29, 476–484. doi: 10.3758/s13423-021-02015-7

Wilson, A. J., and Dehaene, S. (2007). "Number sense and developmental dyscalculia," in *Human Behavior, Learning, and the Developing Brain: Atypical Development*, eds. D. Coch, G. Dawson, and K. Fischer (New York, NY: Guilford Press), 212–238.

Wollman, D. C. (1965). The attainments in English and arithmetic of secondary school pupils with impaired hearing. *Teach. Deaf* 159, 121–129.

Wood, D., Wood, H., and Howarth, P. (1983). Mathematical abilities of deaf schoolleavers. *Br. J. Dev. Psychol.* 1, 67–73. doi: 10.1111/j.2044-835X.1983.tb00544.x

World Health Organization (2021). *World Report on Hearing.* Available online at: https://www.who.int/publications/i/item/world-report-on-hearing (accessed May 27, 2022).

Xing, E., Jordan, M., Russell, S. J., and Ng, A. (2002). Distance metric learning with application to clustering with side-information. *Adv. Neural Inf. Proc. Syst.* 15, 521–528.

Xu, R., and Wunsch, D. (2009). *Clustering*. Hoboken,NJ: John Wiley and Sons.

Zarfaty, Y., Nunes, T., and Bryant, P. (2004). The performance of young deaf children in spatial and temporal number tasks. *J. Deaf Stud. Deaf Educ.* 9, 315–326. doi: 10.1093/deafed/enh034