LUNDS UNIVERSITET

INSTITUTIONEN FÖR PSYKOLOGI

EN: What Eye Tracking Tells Us About Mathematical Difficulties: A Systematic Review

SV: Vad ögonrörelsemätning lär oss om matematiksvårigheter: en systematisk litteraturöversikt

Jonatan Almfjord & Lina Hallberg

Kandidatuppsats HT 2020

Handledare: Petri Kajonius

Examinator: Roger Johansson

Abstract

Eye tracking is a popular method used in mathematics research for investigating underlying cognitive processes. Individuals with mathematical difficulties have problems with mastering mathematical skills and are often in need of extra help. This systematic review aimed to review eleven studies (nine samples) published between 2002 and 2020 to assess whether eye tracking is a useful method for researching mathematical difficulties, and what research using eye tracking can say about mathematical difficulties. Eye tracking did reveal better accounts of cognitive processes compared to verbal reports. The results further showed that individuals with mathematical difficulties often used different strategies and tended to have more and sometimes longer fixations compared with typically developing individuals. In conclusion, eye tracking as a method seemed particularly useful for studying strategy use, discovering why and how errors occur and why response times differ. More research is needed to see longitudinal results, evaluate and compare tools and resources used in educational settings, and make conclusive statements across different cultures and age groups. This field of research would also benefit from clearly presenting all relevant aspects connected to eye tracking.

Keywords: mathematical difficulties, eye tracking, developmental dyscalculia, mathematical learning disorder, systematic review

Sammanfattning

Ögonrörelsemätning är en populär metod inom matematikforskning för att undersöka underliggande kognitiva processer. Individer med matematiksvårigheter har svårt att bemästra matematiska färdigheter och är ofta i behov av extra hjälp. Denna systematiska litteraturöversikt syftade till att granska elva studier (nio urval) publicerade mellan 2002 och 2020 för att ta reda på huruvida ögonrörelsemätning var en användbar metod för att undersöka matematiksvårigheter och vad forskning som använder ögonrörelsemätning säger om matematiksvårigheter. Ögonrörelsemätning redogjorde bättre för kognitiva processer jämfört med muntlig rapportering. Resultaten visade vidare att individer med matematiksvårigheter ofta använder andra lösningsstrategier och tenderar att ha fler och ibland längre fixeringar jämfört med individer med typisk utveckling. Avslutningsvis verkar ögonrörelsemätning vara en användbar metod för att undersöka strategianvändning, för att upptäcka varför och hur fel uppstår och varför responstider skiljer sig åt. Mer forskning behövs för att se longitudinella resultat, utvärdera och jämföra olika hjälpmedel och resurser som används i utbildningsmiljöer. Forskningsområdet skulle också gynnas av att alla relevanta aspekter kopplat till ögonrörelsemätning tydligt presenteras.

Nyckelord: matematiksvårigheter, ögonrörelsemätning, dyskalkyli, inlärningssvårigheter i matematik, systematisk litteraturöversikt

LUNDS UNIVERSITET

Acknowledgment

We would like to thank several people for helping us with this thesis. The researchers Evelyn Kroesbergen, Fae van der Weijden, Anne van Hoogmoed, Sietske van Viersen, Jaccoline van 't Noordende, Riikka-Maija Mononen, and Achim Lilienthal for both helpful feedback and kind words of encouragement through email correspondence. Thank you to Alexia Han for providing the graphic design figures. We would also like to extend our thanks to those of our friends and families who assisted us with proofreading. And finally, a big thank you to our supervisor Petri Kajonius and to our examinator Roger Johansson.

What Eye Tracking Tells Us About Mathematical Difficulties: A Systematic Review

Mathematical abilities are essential in daily routine (Skagerlund et al., 2019), as well as an important topic in school settings and research. Low numeracy skills have been negatively linked to mental and physical health (e.g., Garcia-Retamero et al., 2015; Skagerlund et al., 2019). It also has a negative impact on employment prospects and the economic status of countries (Bynner, 1997; Rivera-Batiz, 1992). Eye tracking has become increasingly popular as a tool to investigate cognitive processes in mathematics research. This systematic review aims to give an overview of the studies conducted using eye tracking to investigate mathematical difficulties. Mathematical difficulties are characterized by low scores on mathematical tests, compared to age-matched peers. By using eye tracking, researchers might get an insight into why this is the case. What studies have been conducted using eye tracking to research mathematical difficulties? Is eye tracking a useful method for studies researching mathematical difficulties, and in what ways have eye tracking data been used to answer critical questions in this area? These are questions we attempt to answer in this review.

Mathematical difficulties are studied in psychology but also in pedagogy, neuroscience, and cognitive science. This makes the area interdisciplinary. Studying mathematical difficulties is relevant in psychology because researching different forms of cognition is important for understanding the human mind. Targeting individuals with mathematical difficulties is especially important since these individuals more often than typically developing individuals suffer from school and work centered anxieties related to mathematics. If cognitive processes of individuals who are thought of as atypical in an area are better understood, it enables researchers to broaden and perhaps question the view of normality. This will also assist in developing new helping tools and adjust learning situations to better suit the individual.

In the following sections, we will start by giving an overview of eye tracking. This includes both theoretical foundations behind its usage, as well as more practical information. Further, we will give an account of mathematical difficulties and similar terms used to target individuals with difficulties in mathematics. Later, different aspects of mathematical abilities, that are tested in the reviewed studies, are explained. Lastly, we will finish the introduction with a brief account of previous research and the aims of the review.

Eye Tracking

Eye tracking is a technique used to gather information about fixation patterns of the human eye. A person's fixations are related to what the person is attending to, and attention is an important component of human cognition. Therefore, knowing what a person is looking at, gives valuable insights into cognitive processes. Eye tracking can be used interactively - as in virtual reality - or in a diagnostic way as in the studies included in this review. When used for a diagnostic purpose the device is tracking the participant's attentional patterns of gaze over the given stimulus. The eye tracking data, both from head-worn and table-mounted devices, is gathered to be analyzed post-trial to see what the participant saw, or missed when completing a task (Duchowski, 2017).

Eye tracking, or eye movement research, has a long history stretching over more than one and a half-century (Wade, 2010). Recent technological advancements have made video-based combined pupil and corneal reflection methods the most commonly used today. The light reflection of the corneal is measured in relation to the pupil which makes the eye tracker able to record the viewer's point of regard. The eye tracker gathers information of the viewer's point of regard relative to the screen of the eye tracker. Eye movements are often partitioned into fixations and saccades (Salvucci & Goldberg, 2000). Fixations are movements of the eyes that stabilize the retina over the object of interest. But fixations are not smooth and stable, rather they are characterized by small movements of the eyes such as tremor, microsaccades, and drift. Fixations are indicators of the location of the viewer's visual attention (Duchowski, 2017). Saccades are quick movements between fixations (Salvucci & Goldberg, 2000). Since the difference between saccades and fixations is not given, researchers need to specify a threshold value in order to identify fixations. There are two main types of algorithms used to identify fixations through the use of threshold values. Dispersion-based algorithms are based on a coordinate system on the X- and Y-axis. Data points close are identified as fixations according to a threshold value, usually around 80 milliseconds. The second type of alternative is velocity-based algorithms. The speed between each sample is identified according to a velocity-threshold. Quicker velocities between samples are identified as saccades and the remaining data is identified as fixations (Duchowski, 2017).

The amount of eye tracking research has increased dramatically as an effect of increased availability and relatively low cost (Duchowski, 2017). As part of the area of cognitive psychology, eye trackers have become increasingly popular for investigating aspects of mathematical difficulties. Now, different concepts related to mathematical difficulties will be explained.

Developmental Dyscalculia and Mathematical Learning Disability

There is no general consensus about the terminology for individuals experiencing difficulties with mathematics in the literature. The terms *developmental dyscalculia*, *mathematical learning disabilities,* and *mathematical difficulties* are used depending on national guidelines and research contexts. All of the aforementioned terms are used in at least one reviewed study and will be defined in the following sections.

Developmental dyscalculia (sometimes referred to as just dyscalculia) and mathematical learning disabilities are often used interchangeably. Both mathematical learning disabilities and developmental dyscalculia are characterized by profound difficulties with numeracy (Kucian & von Aster, 2015; Geary et al., 2007). Developmental dyscalculia and mathematical learning disabilities are commonly identified by using cutoff scores on mathematical tests. The cutoff ranges from <30th percentile to <5th percentile (Geary et al., 2007). The unresolved diagnostic criteria lead to uncertainty of the percentage of individuals with this learning disability (Geary et al., 2007). The suggested prevalence rate is between 3% to 14% depending on the criteria used and country tested (Barbaresi, 2005; Desoete et al., 2004; Gross-Tsur et al., 1996; Koumoula et al., 2004; Lewis et al., 1994; Mazzocco & Myers, 2003; Ramaa & Gowramma, 2002; Shalev & Gross-Tsur 2001). Developmental dyscalculia and mathematical learning disability are shown to persist into adulthood (Shalev et al., 2005; Mejias et al., 2012).

In the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), there are four criteria for diagnosing dyscalculia (American Psychiatric Association, 2013).

1. Difficulties in mastering arithmetic skills, number facts, number sense, calculation, and mathematical reasoning. The difficulties persist for at least six months and do not improve with the provision of an intervention designed to deal with these difficulties.

2. The individual falls substantially and quantifiably below typically developing individuals of the same age. This is assessed through standardized tests with cutoff scores.
3. Dyscalculia may begin in the early years but may also not fully manifest until young adulthood.
4. Before confirming a diagnosis any other disorder or factor such as sensory problems, psychosocial adversity, language problems, or inadequate instructions should be ruled out.

According to the International Statistical Classification of Diseases and Related Health Problems (ICD), developmental learning disorder with impairments in mathematics (also known as mathematical learning disorder), is characterized by the following criteria (World Health Organization, 2019):

1. Severe and persistent difficulties in number sense, accurate calculation, memory of number facts, accurate mathematical reasoning, and fluent calculations.
2. The individual falls markedly below peers in the same chronological or developmental age.
3. The difficulties result in impairments in occupational or academic functioning.
4. Is not due to sensory impairment, disorders of intellectual development, lack of education, neurological disorders, psychosocial adversity, or lack of language skills used in the environmental setting.

The criteria from the DSM-5 and the criteria from the ICD manual (2019) are similar but not identical. Since the diagnostic criteria differ, it is ambiguous whether individuals with certain comorbid conditions would be diagnosed with developmental dyscalculia or mathematical learning disability. Further, both in the ICD manual and in earlier versions of the DSM, developmental dyscalculia has been diagnosed based partly on a discrepancy of IQ and mathematical performance (American Psychiatric Association, 1994; World Health Organization, 2019). This means that the individual's performance on mathematical tests is lower than expected based on their score on an IQ-test. This discrepancy criterion is removed in the DSM-5 (American Psychiatric Association, 2013). Since the DSM-5 was released, individuals who got their diagnosis before 2013 typically fulfill the discrepancy criterion while individuals who got their diagnoses after 2013 will not necessarily showcase this.

Mathematical Difficulties

Mathematical difficulties (sometimes referred to as mathematics difficulties) is a term used in order not to distinguish between individuals who meet the IQ-discrepancy condition and those that do not. It is argued that individuals who meet the discrepancy criteria and individuals that do not, show qualitatively similar cognitive patterns in subitizing, magnitude comparison, and counting (Schindler et al., 2019a). In line with previous research (e.g., Moser Opitz, 2016; Schindler et al., 2020), we choose to use the term *mathematical difficulties* throughout this review, since it includes people with both diagnosed and undiagnosed developmental dyscalculia as well as mathematical learning disabilities. The term has never rested upon the IQ-discrepancy criteria which makes the term more inclusive and possibly less stigmatizing. Individuals with mathematical difficulties are often compared with a group of so-called typically developing individuals. *Typically developing individuals* is a commonly accepted term amongst medical researchers referring to individuals that reach certain age-appropriate milestones. For the present review, the term will refer to individuals who do not have mathematical difficulties.

What types of problems that individuals with mathematical difficulties experience is debated (e.g., Butterworth, 2010; Geary, 2011; Kucian & von Aster, 2015). Suggestions are problems with number sense, problems with subitizing, and problems with number magnitude representation (Moeller et al., 2009; van Hoogmoed et al., n.d.; van 't Noordende et al., 2016). The included studies in this review test these different aspects of mathematical abilities through several different tasks, which will be defined in the following sections.

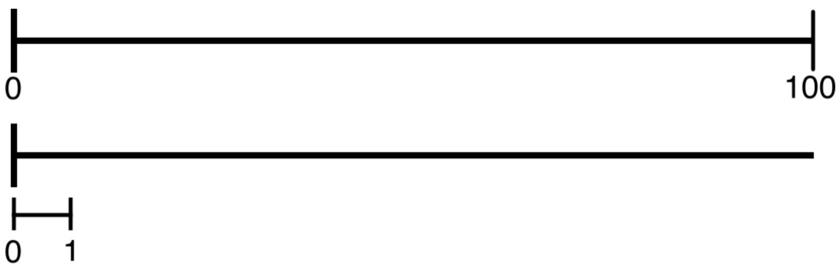
Number Line Estimation

Several studies have suggested that a deficit in number sense might be involved in mathematical difficulties, (e.g., Kucian & von Aster, 2015). Number sense refers to an ability that enables individuals to discriminate between and compare concrete magnitudes approximately. Number sense also includes understanding mathematical concepts such as more than and less than. As part of number sense, number magnitude representation involves the ability to represent magnitudes. ‘Magnitude’ means the relative size of numbers.

On a bounded number line estimation task, the subjects are asked to place a given number on a visual number line, often graded from 1-100 or 1-1000, as close to the target number as possible (Siegler & Booth, 2004), see Figure 1.

Figure 1

Number Line Tasks

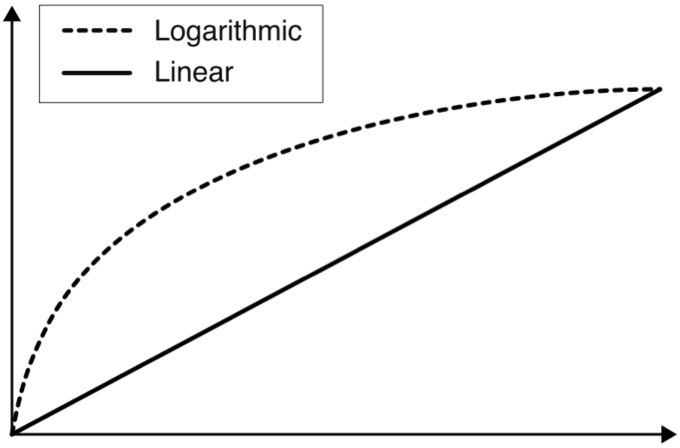


Note. Model of a bounded number line estimation task (top line), with both numbered beginning and ending; and an unbounded number line estimation task (bottom line), which has a numbered beginning and a single measurement unit.

Several studies have found that children with mathematical difficulties have problems placing the requested number near the target position (e.g., Geary et al., 2007; Geary et al., 2008). There are several different models explaining the results on the bounded number line task. A lot of research rests on the assumption that estimations on the bounded number line task reflect a mental representation of a number line. The most popular models to explain these estimations are the logarithmic model and the linear model, see Figure 2. Children's general mathematical ability is significantly related to the shape of their mental number line (Booth & Siegler, 2006). Young children often showcase logarithmic patterns because of their young age and inexperience with larger magnitudes (Booth & Siegler, 2006). Logarithmic patterns mean that numbers closer to zero are valued more, while their deemed value diminish as they approach larger numbers. Older children and adults typically demonstrate estimations that fit a linear model of the mental number line better than a logarithmic model (Booth & Siegler, 2006); Siegler & Opfer, 2003). This means they tend to place the requested number more accurately since the distance between numbers is the same both for smaller and larger quantities.

Figure 2

Logarithmic and Linear Curves



Note. A visual illustration of how a logarithmic curve differs from a linear progression.

Several articles suggest that estimations on the bounded number line are more linked to proportion-judgments than to number magnitude representation on a mental number line (e.g., Barth & Paladino, 2011; Slusser et al., 2013; Sullivan et al., 2011). The researchers applied and compared a proportion-based model with both the linear and the logarithmic models. The proportion-based model, including reference points such as the beginning-, mid-, and endpoint, leads to more accurate predictions. Proportion-based judgments are thought to rely on more advanced numerical abilities. Therefore, the unbounded number line was developed, to give a better and purer measure of number magnitude representation (Cohen & Blanc-Goldhammer, 2011).

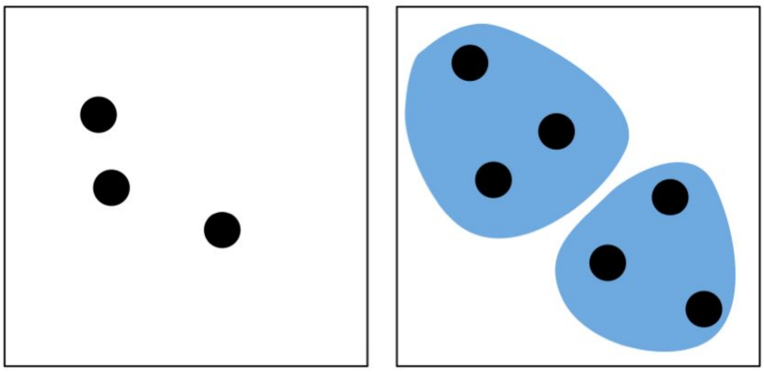
On an unbounded number line, the participants are asked to place a given number on a visual number line (Cohen & Blanc-Goldhammer, 2011). The unbounded number line does only have a beginning point and one scaling unit, usually 0-1 (Cohen & Blanc-Goldhammer, 2011), see Figure 1. Link et al. (2014) suggested that even though the unbounded number line may give a better account of an individual's number magnitude representation skills, the results on an unbounded number task are not related to general arithmetic ability.

Subitizing

The term subitizing was coined in 1949 by Kaufman and colleagues. Subitizing refers to the ability to rapidly, accurately, and confidently report numerosity of small numbers of entities, presented for a short duration of time (Kaufman et al., 1949). This is an automatized perception process that usually makes you able to recognize three to four items at first glance (Mandler & Shebo, 1982; Trick & Pylyshyn, 1994), see Figure 3.

Figure 3

Subitizing and Groupitizing



Note. Illustrations of dots in the subitizing range (left box), and dots in the groupitizing range (right box). The blue shadows show an example of how dots can be divided into smaller subgroups via the process of groupitizing.

Enumeration of entities could also be done through counting (Benoit et al., 2004), or through a similar process called groupitizing (Starkey & McCandliss, 2014). When groupitizing, the set of entities is divided into smaller subgroups which are easier to estimate. Enumeration could also involve recognition of patterns - for example, patterns on a dice - which become familiar and are easily recalled and recognized at a glance, as in subitizing (Ashkenazi et al., 2013). This quick recognition of patterns is sometimes referred to as conceptual subitizing (Schindler et al., 2019a).

Research testing children indicate that the subitizing mechanism in individuals with mathematical difficulties appears to be impaired (Landerl, 2013; Schleifer & Landerl, 2011; van der Sluis et al., 2004). This is suggested since response times for individuals with mathematical difficulties tend to be longer compared to typically developing individuals in subitizing tasks (van der Sluis et al., 2004).

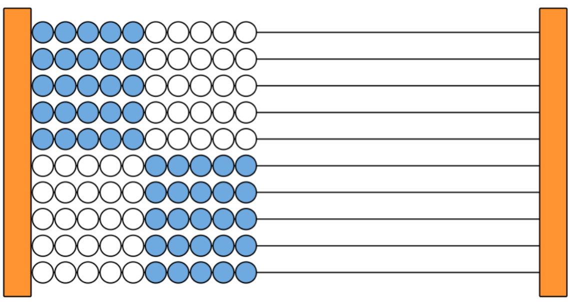
Helping Materials

Another ability, important for mathematics in school-settings, is the ability to perceive partitions of number structures and sets that relate to, for example, the base-ten system (Obersteiner et al., 2014). The base-ten system is often used in classroom settings, for example - as in some of the reviewed studies - with the help of an abacus or the 100-dot-square. These types of helping material are thought to help individuals work with and recognize larger quantities by perceiving partitions into smaller groups (Obersteiner et al., 2014).

A 100 bead-abacus has ten poles with ten beads on each pole, see Figure 4. Each pole contains five white beads and five beads of a different color. The first five poles have a reversed color order compared to the last five rows. This represents the 25- and 50-structure (Schindler et al., 2019a). The 100-dot-square is constituted by 100 same-colored dots in ten rows of ten dots each, see Figure 5. Between the dots in column five and six and between row five and six, there is a bigger gap representing the 25- and 50-structure (Schindler et al., 2019a). Another type of helping material is dice-like patterns, see Figure 6. When using dices, individuals often learn to recognize the patterns since they are arranged canonically. This means that one could grasp the magnitude quickly instead of counting the dots one by one.

Figure 4

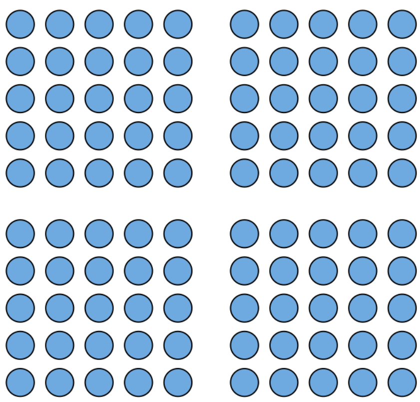
100 Bead-Abacus



Note. Illustration of a 100 bead-abacus.

Figure 5

100-Dot-Square



Note. Illustration of a 100-dot-square.

Figure 6

Dice-Like Patterns

Note. Illustration of dice-like patterns.

Previous Research

Eye tracking as a method has become increasingly popular in mathematics research over time (Lilienthal & Schindler, 2019). Several studies have been conducted focusing on typically developing individuals rather than individuals with mathematical difficulties (e.g., Heine et al., 2009; Schneider et al., 2008; Sullivan et al., 2011). Five review studies have summarized studies using eye tracking equipment in the field of education and four of them mainly focus on mathematics (Lai et al., 2013; Lilienthal & Schindler, 2019; Mock et al., 2016; Perttula, 2017; Strohmaier et al., 2020). Strohmaier et al. (2020) included a subheading reviewing eye tracking and mathematics for individuals with learning difficulties. This subheading includes five studies included in this review as well as three articles focusing on autism (Winoto et al., 2017), Down syndrome (Abreu-Mendoza & Arias-Trejo, 2015), and developmental coordination disorder (Gomez et al., 2017). These last three articles will not figure in this review. We chose to focus on studies with the main focus of mathematical difficulties and therefore chose to exclude studies where the main focus was diagnoses or disabilities other than mathematical difficulties. This also follows the diagnostic criteria from the ICD manual (2019) which excludes individuals with neurological disorders and intellectual disorders from getting the diagnosis of mathematical difficulties. Apart from this, we found three studies from 2002, 2015, and 2019 not included in Strohmaier et al. (2020). We also found one article published after Strohmaier et al. (2020) was published. Finally, this review also includes results from one unpublished study.

The review of Strohmaier et al. (2020) covered over 100 articles. The numerosity of studies included in the review affected the depth of analysis in the subfield of eye tracking related to mathematical difficulties. Therefore we wanted to make the results in this area clearer. Individuals with mathematical difficulties face anxieties and problems in school and work settings, which give reasons to focus attention on these individuals who have the potential to benefit the most from intervention and research. Some additional research has been conducted in recent years, which makes a review relevant to get an overview of the field.

Purpose

The present study aims to provide the body of scientific literature with an up-to-date systematic review on eye tracking as a method of investigating mathematical difficulties.

The central research questions are:

1. What studies have been conducted on the topic of eye tracking as a method in researching mathematical difficulties?
2. Is eye tracking a useful method for studies researching mathematical difficulties?
3. In what ways have eye tracking data been used to answer critical questions in mathematical difficulties research?

Method

The present study consists of a systematic review of research on the topic of eye tracking as a method of investigation of mathematical difficulties. As guidelines for the review, the most recent version of the PRISMA Checklist (Moher et al., 2009) was used.

Search Strategies

A three-step process was used to identify relevant studies. First, a systematic database search was conducted in PubMed, PsycNet, and LUBsearch (including Directory of Open Access Journals, MEDLINE, ERIC, Academic Search Complete, and ScienceDirect) with the inclusion and exclusion criteria listed below. The following search string was used, referring to abstracts and titles:

eye tracking AND dyscalculia OR mathematical learning disabilit* OR mathematical learning difficult* OR mathematical disabilit* OR mathematical difficult*

Second, we carefully checked the references of the studies found in the first step for any additional material to be found there, in which case they were added to the review. Third and finally, through email, we got in touch with a number of the most frequently occurring researchers in the included studies and asked if they knew of any unpublished studies. This step was of particular importance in order to reduce the risk of bias across studies.

Inclusion Criteria

To qualify for the present systematic review, studies had to align with the following criteria:

(a) The study included at least one empirical study.
(b) The study included eye tracking as a method of investigating mathematical difficulties.

It was decided that no inclusion criteria were needed for region of origin or year of publication.

Exclusion Criterion

Studies that aligned with the following criterion were after consideration excluded from the review:

(a) The focus of the study was disabilities or diagnoses other than mathematical difficulties.

Search Result

The three-step search process described gave the following outcome. The systematic database search was conducted in mid-December 2020 and identified seven studies. Checking for additional studies through the reference lists, resulted in three additional studies. Reaching out to researchers in the field provided us with one unpublished paper, which we were allowed to include in our review.[1] This left us with a total number of eleven included studies.

[1] The mentioned paper, van Hoogmoed et al. (n.d.), is an unpublished study that is highly similar to van Wijk (2017). Both studies will hence figure in the references section.

Ethics

Considering that the present review exclusively treated data from previously conducted studies that were already ethically reviewed, no additional ethical considerations were accounted for.

Results

Included Studies

Eleven studies were included and are presented in Table 1, ordered alphabetically after the first author. Since all studies included a control group with typically developing individuals of at least the same number as individuals with mathematical difficulties, the number of participants and ages of both groups are presented separately when possible. Furthermore, the country of origin for the study, eye tracking equipment used, and what test was being used, are presented.

The studies Schindler et al. (2019a) and Schindler and Lilienthal (2018) were based on the same participant sample but analyzed the results differently. The studies Schot et al. (2015) and van Viersen (2013) were almost based on the same sample, since all participants were the same except for one additional child with mathematical difficulties that were added in Schot et al. (2015). In the Results, we will refer to all studies to support a statement even if they are based on the same sample. In Table 2 this is clarified, see Appendix. There, studies based on the same data are shown with square brackets.

Table 1

Studies Included in the Systematic Review

#	Study	Participants				Ctry.	Eye Tracking Device	Test Used
		No. MD	No. TD	Age MD [M(SD)]	Age TD [M(SD)]			
1.	van Hoogmoed et al. (n.d.)	15	15	10.60 (1.16)	10.40 (0.85)	NLD	Tobii T60	Unbounded NLT and cued retrospective reporting.
2.	Moeller et al. (2009)	2	8	10.38 (0.29)	10.42 (0.25)	AUT	EyeLink 1000	Name the number of black dots within a white square.
3.	van 't Noordende et al. (2016)	14	14	11.09 (1.10)	10.71 (0.89)	NLD	Tobii T60	0-100 NLT and 0-1000 NLT.
4.	Rottman & Schipper (2002)	4	4	-	-	DEU	-	Addition and subtraction with the help of 100-dot-square.
5.	Schindler et al. (2019a)	10	10	10.83 (0.50)	10.58 (0.42)	DEU	Tobii Pro Glasses 2	Estimate numbers on a 100-beads abacus and 100-dot-square.
6.	Schindler et al. (2019b)	69	59	10.75 (0.58)	10.75 (0.58)	DEU	Tobii X3-120	Estimate numbers on a 100-beads abacus and 100-dot-square.
7.	Schindler et al. (2020)	10	10	10.83 (0.50)	10.58 (0.42)	DEU	Tobii Pro Glasses 2	Name number of dots (1-9) showing on a screen.
8.	Schindler & Lilienthal (2018)	10	10	10.83 (0.50)	10.58 (0.42)	DEU	Tobii Pro Glasses 2	Estimate numbers on a 100-beads abacus and 100-dot-square.
9.	Schot et al. (2015)	2	10	9.78 (0.04)	9.10 (0.60)	NLD	Tobii T60	Bounded NLT.
10.	van Viersen et al. (2013)	1	9	9.00 (-)	9.10 (0.60)	NLD	Tobii T60	Bounded NLT in both symbolic and non-symbolic versions.
11.	van der Weijden et al. (2018)	8	8	22.85 (3.85)	22.05 (2.66)	NLD	Tobii T60	Bounded and bounded NLT.

Note. Cat. = Category. No. MD = Number of participants with mathematical difficulties. No. TD = Number of typically developing participants. Age MD = Age of participants with mathematical difficulties. Age TD = Age of typically developing participants. Ctry. = Country. NLD = Netherlands; AUT = Austria; DEU = Germany (International Organization for Standardization, 2020). NLT = Number line task.

Eye Tracking as a Method to Investigate Mathematical Difficulties

Eye Tracking Equipment and Technical Aspects

The studies used different threshold values for what they counted as a fixation. Several studies did not have or did not report, a threshold value. Others reported using standardized values such as 83 ms, which is the standard-setting on some eye trackers. Not reporting information about threshold values makes replication difficult. What counts as fixations and how the fixations are interpreted are important for analyzing the results of the studies. If interpreters use different threshold values to analyze the same data, the results will differ. Using different threshold values also makes comparisons between studies difficult (Salvucci & Goldberg, 2000). Generally, fixations are interpreted as a measure of what the individual is attending. But, sometimes, as in Rottmann and Schipper (2002), several seconds long gazes on one area of the material were reported. These gazes were analyzed, not as representing what the individual was attending to, but rather that the individual counted in the head while staring at one set point. Even if the analysis of eye tracking data is somewhat qualitative, it would be useful to differentiate between attentional fixations and non-attentional fixations. This would preferably be done and presented before the studies and not in the interpretation of the results.

Schot et al. (2015) and van Viersen et al. (2013) were more or less based on the same sample but analyzed the eye tracking data differently. Effects were identified as smaller when using quantitative analysis. Several different eye tracking devices were used in the studies. Tobii T60 (table-mounted device) was used in five studies, Tobii Pro Glasses 2 (head-worn device) in three, Tobii X3-120 (table-mounted device) in one, and EyeLink 100 (table-mounted device) in one. In one study (Rottmann & Schipper, 2002), it was unclear what device was used. It is important to give an account of the equipment used for transparency and replicability. Aiming for a transparent presentation of all relevant eye tracking-aspects would decrease the risk of reporting non-significant results and increase the validity and reliability of the findings.

Is Eye Tracking Useful?

Giving a verbal report of the strategies used when doing an arithmetic calculation has the advantage of accessibility. No extra equipment or analysis is needed in most cases, which makes it affordable and practical. Schindler and Lilienthal (2018) compared the reported strategy uses gathered with verbal reports versus eye tracking. In most cases, eye tracking gave more information than verbal reports. This was especially true for students with mathematical difficulties.

Even if the participants correctly expressed the strategies used verbally, eye tracking often provided more detailed information of when and where these strategies were used. Eye tracking gave clues to why processing times did differ and made researchers able to detect where and when the participants made mistakes. Eye tracking gave better insights into the cognitive strategies than verbal reports. This knowledge about students' strengths and weaknesses could lead to better-adjusted learning situations to suit the individual. Some possible explanations are proposed in the studies. These are anxieties connected to mathematics, inexperience in expressing strategy thoughts, shyness, poor language skills, forgetting or denying some strategy use because of social demands, possibly affected by follow-up questions from the researcher on the verbal report which guided the students to certain norms (Schindler & Lilienthal, 2018).

Rottmann and Schipper (2002) discovered that children who rank high in mathematics by their teacher often do not use helping material to solve addition and subtraction problems, while children who are ranked low in mathematics by their teacher tend to use the helping material often - but do so incorrectly, which leads to errors. They also showed that for high achievers in mathematics, the helping material was used to demonstrate their strategies in verbal reports afterward, rather than actively using it in their own arithmetic processes. This showed that eye tracking can give information, perhaps unknown even to the participants in the experiments, in a way that verbal reports often do not.

Eye tracking seems to be useful, especially for detecting different forms of strategies. What types of strategies an individual uses is connected to whether they have mathematical difficulties or not. This opens up the possibility to use eye tracking as a diagnostic tool to complement available tests to detect mathematical difficulties.

Eye Tracking as a Diagnostic Tool

Several studies have suggested that eye tracking could be used as a diagnostic tool when detecting mathematical difficulties. Schindler et al. (2020) proposed that eye tracking has diagnostic potential for detecting individuals with mathematical difficulties when used to study enumeration processes of small sets of items. Results from Schot et al. (2015), van 't

Noordende et al. (2016), and van Viersen et al. (2013) suggest that using eye tracking in combination with bounded number line task could be beneficial as a complement to other diagnostic measures. But the results from van der Weijden et al. (2018) who tested adults with and without dyscalculia on the bounded number line task, found highly similar results between the groups. This indicates that eye tracking in combination with bounded number line task is more efficient as a potential diagnostic tool when investigating children rather than adults. Using eye tracking as a diagnostic tool combined with the unbounded number line task could be effective since there are group differences in estimation strategies. However, the accuracy on the unbounded number line task does not differ much between individuals with mathematical difficulties and typically developing individuals (van der Weijden et al., 2018; van Hoogmoed et al., n.d.), indicating that it, perhaps, would be more interesting to test individuals on other tests. Overall, eye tracking has the potential to be useful as a diagnostic tool.

Eye Tracking Data in Mathematical Difficulties Research

Subitizing and Groupitizing

Moeller et al. (2009), Schindler and Lilienthal (2018), Schindler et al. (2019a), and Schindler et al. (2020), studied subitizing and groupitizing in children. Results from these studies indicate that the subitizing ability is impaired in children with mathematical difficulties. When analyzing eye tracking data Moeller et al. (2009), Schindler and Lilienthal (2018), Schindler et al. (2019a), and Schindler et al. (2020) discovered that children with mathematical difficulties used different strategies compared with typically developing children. Children with mathematical difficulties demonstrated a gaze behavior where they appeared to look at all the dots (instead of subitizing) more often compared with typically developing children. Sometimes this is showcased in more fixations but could also demonstrate in longer fixation duration at the objects presented.

Schindler et al. (2020), who also tested groupitizing, saw that students with mathematical difficulties used groupitizing significantly less often compared with typically developing students. The eye tracking-data from the children with mathematical difficulties suggests that they looked at all dots, one by one, more often. This could reflect counting all dots one by one. This falls in line with previous findings from Fischer et al. (2008), Landerl (2013), Schleifer and Landerl (2011), and van der Sluis et al. (2004), who all detected longer response times for students with mathematical difficulties. These results also answer the question of why children with mathematical difficulties tend to have longer response times compared with typically developing children. Children with mathematical difficulties are in many cases looking at all entities, even in small quantities, one by one. This gives some evidence to the notion of a qualitative difference between children with mathematical difficulties and typically developing children. This implies that one could give children with mathematical difficulties extra time to solve subitizing problems with their own strategies.

Unbounded Number Line Task

Van der Weijden et al. (2018) and van Hoogmoed et al. (n.d.) tested children and adults on the unbounded number line task. The eye tracking data from van der Weijden et al. (2018) suggests that typically developing individuals used dead-reckoning more often than individuals with mathematical difficulties. Dead-reckoning is a strategy characterized by using one scaling unit and a continuous adding of this scaling unit until the target position is reached. The eye tracking data from van der Weijden et al. (2018) and Hoogmoed et al. (n.d.). suggests that individuals with mathematical difficulties did instead use direct estimations more often than typically developing individuals. This strategy is characterized by first looking at the unit and then determining where it approximately should be (van der Weijden et al., 2018). These results do not show any group-differences in accuracy, indicating that direct guesses might not be a bad method in this case. If the unbounded number line is an accurate measure of number magnitude sense, results from these studies indicate that this is not impaired in individuals with mathematical difficulties. These results also show that unbounded number line tasks are difficult both for typically developing individuals and for individuals with mathematical difficulties. Since there do not seem to be any group differences in accuracy on the unbounded number line task, and since Link et al. (2014) suggests that the results on the unbounded number line task are not related to general arithmetic ability, we propose that it would be more interesting to test individuals with mathematical difficulties on other tests.

Bounded Number Line Task

Schot et al. (2015), van der Weijden et al. (2018), van 't Noordende et al. (2016), and van Viersen et al. (2013) tested children and adults on the bounded number line task. They tested group differences in estimation strategy and whether the strategy used was functional. Predefined strategies included the counting up and counting down strategy where the beginning or the endpoint was used as a reference, and the midpoint strategy when the midpoint was used as a reference point. Another estimation strategy is called estimating the small unit, for which the participant divides the target number into two parts and places the major part rather carefully on the line while the smaller unit is estimated loosely. Unidentifiable strategies are strategies that either are a mix of different strategies or do not align with any predefined strategy. Strategies were also classified as functional or dysfunctional. A functional strategy uses a reference point close to the target number. A dysfunctional strategy is characterized by the use of reference points far away from the target number.

The eye tracking data from van 't Noordende et al. (2016) found multiple non-significant differences between the experimental groups with regards to estimation strategies. These differences consisted of less usage of reference points in the group of individuals with mathematical difficulties compared with typically developing individuals. And this difference was especially large with regards to using the counting up strategy. However, individuals with mathematical difficulties more often than typically developing individuals used a combination of both midpoint and endpoint, all three reference points simultaneously, and guessing. The data from van 't Noordende et al. (2016) also showed that individuals with mathematical difficulties used less functional strategies than typically developing individuals. Schot et al. (2015) and van Viersen et al. (2013) showed that children with mathematical difficulties tended to have more unidentifiable strategies compared with typically developing children. They also showcased more dysfunctional strategy use where the fixations often were centered around the midpoint even for very high or very low numbers.

Eye tracking data from van der Weijden et al. (2018) revealed that adults with mathematical difficulties did not differ notably in strategy use compared with typically developing individuals. The only difference is that individuals with mathematical difficulties less often than the control group checked whether they had placed the target number correctly and more often than typically developing individuals estimated the small unit. However, the group of individuals with mathematical difficulties did showcase less functional reference point combinations when the target was a number close to the midpoint. Another finding from the data was that in general, individuals with mathematical difficulties tended to use reference points below the target number, from which they counted up, while typically developing individuals sometimes used reference points above the target, from which they counted down. The results from these four studies indicate that the bounded number line task is more difficult for children with mathematical difficulties than for typically developing children. As adults, individuals with mathematical difficulties and typically developing individuals showcase results that were highly similar both in accuracy and strategy use. This indicates that proportion-based judgments may take longer time to grasp for children with mathematical difficulties, but that this skill probably will be mastered.

Furthermore, the bounded number line has long been theorized to be a measure of a mental number line. The shape of this number line was thought to range from logarithmic to linear depending upon age and mathematical abilities. Recently the bounded number line task has received some criticism since the task might reflect proportion-based judgments, thought to be a higher cognitive skill, rather than pure number magnitude representation on a mental number line. Indeed, the eye tracking research based on the bounded number line confirms the criticism. The eye tracking data, especially for successful estimations, is characterized by a large proportion of fixations focused around reference points on the line. This suggests that reference points are important for successful estimations.

Tests using Commonly used Helping Material

Rottmann and Schipper (2002), Schindler and Lilienthal (2018), Schindler et al. (2019a), Schindler et al. (2019b), and Schindler et al. (2020) used eye tracking to study eye movements when the participants were working with common learning material used in school settings. These studies indicate that children with mathematical difficulties tend to use these materials to count dots one by one instead of using them as their typically developing peers who tend to use structures in the material (rows on an abacus or fields and rows on a 100-dot-square). Rottmann and Schipper (2002) showed that younger children with mathematical difficulties used the learning material more often than typically developing children. Children with mathematical difficulties tended to use the learning material more often compared with the control group. But the eye tracking-data suggests that often the materials are used incorrectly since many fixations indicated that children were looking at rows and dots other than the target number. Several children looked at the first dot in the row when asked for a whole-ten number, such as 30, when instead, 30 is the last dot on the row. By using eye tracking, the researchers could discover how the children used the materials and therefore, when and how an error occurred. Schindler and Lilienthal (2018) and Schindler et al. (2019a) tested slightly older children on the same task and found no differences in error rates between the two tested groups but the response time for children with mathematical difficulties was longer which probably reflects a tendency for these children to use more basic strategies. Children with mathematical difficulties tended to prefer counting one by one instead of using structures in the material and also preferred addition over subtraction. Schindler et al. (2020), who tested viewing and naming of canonical arrangements of dice-like patterns, suggested that students with mathematical difficulties may have problems with implicit pattern recognition or conceptual subitizing. This is in line with the results from Ashkenazi et al. (2013). Students with mathematical difficulties looked at all the dots significantly more often when viewing and naming dice-like patterns, compared with typically developing students. These results raise the question of whether these types of helping materials are beneficial for students with mathematical difficulties since children with mathematical difficulties showcase problems with using and perceiving structures and patterns in these materials. This might lead to questions about to what extent these types of helping materials should be used, since their effectiveness can be questioned.

Focal Population

The number of participants was generally low. Apart from Schindler (2019b) who used over 100 participants, the number of individuals with mathematical difficulties were no larger than 14. Large samples are needed to make conclusive judgments of the results. The age of participants across studies varied only slightly. Most of the studies, except van der Weijden et al. (2018; which had a mean of 22.85 for the participants with mathematical difficulties), and Rottmann and Schipper (2002; which did not specify more than that "German second graders" participated), included participants in the range between 8 and 12 years old.

The exact definition of what counts as mathematical difficulties is not clear and may vary between countries. This makes sampling dependent on country-specific conventions to some extent since the cutoff scores and also the diagnostic tests per se, differ. In all studies except one, individuals judged as having mathematical difficulties had first undergone some form of testing of mathematical abilities and in some cases complementary interviews before participating in the experiments. Several participants were also identified using the IQ-discrepancy criteria. Rottmann and Schipper (2002) used teacher evaluation to select four children who performed among the best in the class and four children who performed among the worst. Using teacher ratings, instead of tests, leads to an increased risk of bias. This method of selecting participants is not optimal since the representability of both groups is questionable. Just because someone is judged as likely to perform poorly on arithmetic tests, does not mean that this person could be diagnosed with mathematical difficulties, and the children that are judged as likely to be among the best-performing, are not representative of typically developing children. Designing an experiment where the results from low-achievers are compared with high-achievers results in greater group-differences. Designing an experiment where the results from low-achievers are compared with high-achievers results in greater group-differences but might not be suitable in order to draw conclusions about individuals with mathematical difficulties.

Age as Factor

The age of the participants in the included studies spans from 8-year-olds to adults. Discrepancies between typically developing individuals and individuals with mathematical difficulties seem to relate somewhat to the age of the participants tested. Greater group differences are discovered testing younger children compared to older children or adults. This is, to our best understanding, a novel finding. A possible explanation for this might relate to familiarity with the material tested. Younger children are not as used to mathematical education and mathematical testing as older individuals. Younger children are not as used to working with larger quantities which makes it more difficult to relate to them (Ashcraft & Christy, 1995; Hamann & Ashcraft, 1986). Another possibility is that the mathematical difficulties an individual faces tend to reduce with age. It would be interesting to study what types of problems related to mathematical difficulties in children that stay with the individual until adulthood and what problems that are overcome.

Discussion

The present study aimed to provide the body of scientific literature with an up-to-date systematic review on eye tracking as a method of investigation of mathematical difficulties. The research questions included:

1. What studies have been conducted on the topic of eye tracking as a method in researching mathematical difficulties?
2. Is eye tracking a useful method for studies researching mathematical difficulties?
3. In what ways have eye tracking data been used to answer critical questions in researching mathematical difficulties?

In pursuing this aim, empirical studies covering the topic were collected, summarized, and compared. Eleven studies and nine samples were included in the review. To summarize the conclusions, the top-findings include the following:

- Eye tracking gives more detailed and accurate accounts of strategy use compared with verbal reports, and could be of diagnostic value for detecting mathematical difficulties when used while working with helping materials, the bounded number line task, or subitizing. Therefore, eye tracking is likely a useful method for researching mathematical difficulties.
- The reviewed studies used eye tracking in order to answer questions about subitizing and groupitizing, about strategy use on the bounded and unbounded number line task, and common helping materials. Children with mathematical difficulties have difficulties with subitizing and groupitizing and showcase differences in strategy use on the bounded number line task compared to typically developing children. They also tend to use helping materials to look at all items one by one rather than using patterns in the materials.

Is Eye Tracking Useful?

Eye tracking seems to give more detailed information about strategy use when used with individuals with mathematical difficulties, compared with verbal reports. It reveals that students with mathematical difficulties tend to have more and sometimes longer fixations, which explains longer response times. With the help of eye tracking, researchers and teachers can also discover why and how students make errors. Why and how an error occurs is usually information that the students are not aware of themselves. Knowing why, is often the first step towards understanding for students. Knowing about the students' cognitive processes are also relevant for teachers in order to provide helpful explanations to the pupils. Since children with mathematical difficulties tend to showcase more fixations and differences in strategy use, eye tracking could be used as a diagnostic tool. Hence, eye tracking is likely a useful method for researching mathematical difficulties. When more research has been conducted, patterns could potentially emerge which opens up for an automated computerized diagnostic process. This field of research would also benefit from presenting all relevant aspects connected to eye tracking. This will be discussed even more under Further Research.

Eye Tracking in Mathematical Difficulties Research

Several aspects of the problems that individuals with mathematical difficulties face have been studied with the help of eye tracking. These include subitizing, groupitizing, the bounded and unbounded number line task, using commonly used helping materials, and pattern recognition. The results from the included studies support previous research in suggesting that the subitizing ability is impaired in individuals with mathematical difficulties. More research is needed, both testing individuals outside of Europe but also to test individuals of different ages, since subitizing tasks have only been tested with children. There could be a qualitative difference between individuals with mathematical difficulties and typically developing individuals in the abilities to subitize, groupitize, and recognize patterns. Also, with the knowledge that the subitizing ability possibly is impaired in children with mathematical difficulties, the learning situation could be adjusted. If it is difficult to use subitizing, groupitizing, and pattern recognition, maybe individuals with mathematical difficulties could be handed more time to solve problems with other strategies? Children with mathematical difficulties also seem to have problems grasping structures in helping materials. Since helping materials often are used with students who need extra help, it is important to investigate whether these resources foster mathematical skills in individuals with mathematical difficulties or not.

Several studies have also seen that individuals with mathematical difficulties use other strategies when doing the unbounded and the bounded number line tasks. More research could potentially improve pedagogic methods used in education to better align with the strategies that individuals with mathematical difficulties prefer. The study testing adults on the bounded and unbounded number line tasks saw no differences between the groups in accuracy. This raises questions of whether mathematical difficulties are reduced with age. We did see a tendency for greater group differences when younger participants are included.

Methodological Considerations and Limitations

All of the reviewed studies, except Schindler et al. (2019b), had few participants. While this weakness cannot be mitigated, it was one of the reasons to conduct the present review. Individually, the studies are weak because of the low number of participants. But when the results are summed up and analyzed, patterns begin to emerge. Still, more research is needed to test larger quantities of individuals. To increase generalizability, studies testing individuals outside of Europe are required. The results from the review are not conclusive and this leaves a lot of room for more research.

One manuscript included in this review, Hoogmoed et al. (n.d.), is unpublished. Since this study has not been peer-reviewed, a valuable instance of quality control is lacking. Even if this is a risk that might affect the quality of the conclusions in the reviewed studies and therefore, the quality of this review, we chose to include it since the number of published studies was relatively small.

One could argue that a limitation of the present review is that it excluded studies focusing on individuals with autism, Down syndrome, and developmental coordination disorder. According to the ICD manual (2019), individuals with these conditions would not be diagnosed with mathematical difficulties since these conditions are classified as neurological disorders as well as intellectual disabilities. We chose to use the diagnostic criteria from this manual since it is widely used, but we remain unsure if it was the right choice. Since many individuals with autism, Down syndrome, and developmental coordination disorder showcase similar difficulties with mathematics, one could argue that these individuals do indeed have mathematical difficulties. Just as the diagnosis of dyscalculia until recently rested on the IQ-discrepancy criteria in the DSM-5 (2013), perhaps mathematical difficulties rest on the assumption of exclusion of individuals with neurological disorders and intellectual disabilities.

A limitation affecting the clarity in this review concerns the discrepancy between the number of samples and the number of studies. Since several studies build upon empirical data published in another study, there are more studies than samples. To make this clearer, we stated this at the beginning of the Results. We also created Table 2, see Appendix, in which square brackets are used when studies are using the same or about the same data. Still, there might be misinterpretations concerning the references in the text since we chose to refer to all studies, even if they were based on the same data. We wanted to add all information we could find to this review and therefore, did not exclude studies that were built on the same empirical findings since the results differed. The downside to this is that sometimes several studies are cited that were based on the same sample.

On the topic of studies reviewed, we would like to raise awareness of methodological considerations in Rottmann and Schipper (2002) that might affect the results of this review. Rottmann and Schipper (2002) chose groups that consisted of four children who tended to perform poorly on arithmetic tests and four children who tended to perform among the best on arithmetic tests. It is not clear if the children participating in the group testing mathematical difficulties had mathematical difficulties according to the ICD manual (2019) and DSM-5 (2013). If this was the case the study would, according to our inclusion criteria, not partake in this review.

There are no exact definition of mathematical difficulties and a common way of testing is by using cutoff scores on standardized tests. The unresolved diagnostic criteria are a problem for the reliability and the validity of the results. In Schindler et al. (2019b) the proportion of individuals analyzed as having mathematical difficulties was higher than the individuals that were analyzed as typically developing. The sample was drawn from a single school in Germany. Of course, there is a possibility that this school's students perform particularly poorly on mathematical tests and that the proportions of students with mathematical difficulties are very high. Yet, it seems unlikely that this is the case. If more than 50% of students are considered below average, maybe, there might be something wrong with the test.

Another problem concerns the term typically developing individuals. The term is frequently used in the literature but is never properly defined. Do typically developing individuals include individuals who tend to perform exceptionally well on arithmetical tests or are these top-performing individuals equally considered as atypical? Is it not reasonable to exclude the best performing individuals in the same way as individuals who perform poorly are excluded from the typically developing? A proper definition of typically developing is needed in order not to exclude people on arbitrary grounds.

Some studies have either vague or missing reports of the type of eye tracking equipment used, threshold values of what counts as fixations, and clear suggestions of how the fixations should be analyzed. This is a limitation. The reports of threshold values could be clarified and could benefit from conventions in the field which will make comparisons easier. Under the subheading of Further Research, a couple of suggestions of how to make the theoretical foundations clearer are presented. We can, however, identify a trend over time in the studies leaning towards more detailed and clear information about the eye tracking equipment used, which is promising since this increases the possibility to replicate studies.

When conducting database searches there is always a risk of not detecting all the relevant studies. We have tried to minimize this risk by emailing frequently published researchers and asking for unpublished material and studies we could not find by ourselves. We got several replies and valuable feedback which we are grateful for. But, there is always the risk of missing out on relevant material. Finally, two studies included in the review were written in Dutch and German and translated with Google Translate. Reading a text through Google Translate is not optimal and this might have affected our results.

Further Research

For further research we have a couple of recommendations. First, the need for more longitudinal studies is clear. None of the included studies measured the participants on more than one occasion. Following the development of individuals with mathematical difficulties over time would provide valuable data on how they progress throughout their lifespans. The results from van der Weijden et al. (2018), which examined adults on the unbounded number line task, raise questions about whether mathematical difficulties are reduced with age. On the other hand, results indicate that the subitizing mechanism is impaired in children with mathematical difficulties. Do older individuals develop this mechanism or are this rather a feature more specific for individuals with mathematical difficulties? Do they simply need more time to look at all the entities one by one? This needs to be addressed in further research.

Also, to support individuals with mathematical difficulties, it would be helpful to conduct extensive eye tracking studies focusing on evaluating and comparing different tools and resources used in educational settings. One study in the present review concluded that although the examined tool was used broadly in school, the eye tracking data suggested that it was not very helpful for the students. It does not seem unlikely that teachers might assume that the available tool does help the student, while it does in fact not.

Finally, information about the equipment and settings used needs to be presented to the reader. Inspired by suggestions from Strohmaier et al. (2020), all relevant aspects connected to eye tracking needs to be clarified. Are all fixations, independent of the duration, position, and when they are made, given equal weight in the analysis? Fourth, in line with Strohmaier et al. (2020), the area in most desperate need of clarification is what eye patterns say about cognitive processes. What do fixations mean? Do all researchers assume all gazes to be attentive, and what signals a non-attentive fixation? Lastly, these assumptions should be presented to the reader before the interpretation of results.

Conclusions

The present study aimed to provide an up-to-date systematic review on eye tracking as a method of investigating mathematical difficulties. Empirical studies covering the topic were collected, summarized, and compared. Eleven studies and nine samples were included in the review. The results indicate that eye tracking generally gives more information than verbal reports and has diagnostic potential for detecting mathematical difficulties when used subitizing, using common helping materials, and doing the bounded number line task. Eye tracking is effective for detecting differences in strategy use and why and how errors occur and response times differ. Some critical questions about mathematical difficulties have been studied with the help of eye tracking. This research indicates that the subitizing and the groupitizing mechanisms seem to be impaired in children with mathematical difficulties. Children with mathematical difficulties tend to count small quantities one by one. This is in line with previous research. Up until recently, the bounded number line task has been used to measure number magnitude representation. The eye tracking research included in this review supports the suggestion that the bounded number line task rather measures proportion-based judgments. This review suggests that children with mathematical difficulties tend to have difficulties with proportion-based judgments on the bounded number line task. On other tasks, such as the unbounded number line task and using helping materials, individuals with mathematical difficulties tend to use different strategies and have longer response times compared with typically developing individuals. The differences between groups are somewhat related to the age of the participants. Younger participants tended to generate larger group differences compared with older participants. This is, to the best of our knowledge, a novel finding.

The current studies and this systematic review are important since they can provide a foundation from which further research can be made - research that in turn can lead to a better understanding for individuals with mathematical difficulties and better mathematics education in general. However, the samples in the studies are relatively small and questions remain about the unclear presentation of all relevant aspects of the eye tracking data and analysis. More research is needed to make conclusive statements about strategy use across different cultures and age groups.

Funding and Conflict of Interest Statement

The present review was conducted as part of a bachelor's thesis and no funding was received. The authors declare no conflict of interest.

References

References marked with an asterisk (*) indicate studies included in the present review.

Abreu-Mendoza, R. A., & Arias-Trejo, N. (2015). Numerical and area comparison abilities in Down syndrome. *Research in Developmental Disabilities*, *41-42*, 58-65. https://doi.org/10.1016/j.ridd.2015.05.008

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.).

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Ashcraft, M., & Christy, K. (1995). The Frequency of Arithmetic Facts in Elementary Texts: Addition and Multiplication in Grades 1-6. *Journal for Research in Mathematics Education*, *26*(5), 396-421. https://doi.org/10.2307/749430

Ashkenazi, S., Mark-Zigdon, N., & Henik, A. (2013). Do subitizing deficits in developmental dyscalculia involve pattern recognition weakness? *Developmental Science*, *16*(1), 35-46. https://doi.org/10.1111/j.1467-7687.2012.01190.x

Barbaresi, W. J., Katusic, S. K., Colligan, R. C., Weaver, A. L., & Jacobsen, S. J. (2005). Math Learning Disorder: Incidence in a Population-Based Birth Cohort, 1976-82, Rochester, Minn. *Ambulatory Pediatrics*, *5*(5), 281-289. https://doi.org/10.1367/A04-209R.1

Barth, H. C., & Paladino, A. M. (2011). The development of numerical estimation: Evidence against a representational shift. *Developmental Science*, *14*(1), 125-135. https://doi.org/10.1111/j.1467-7687.2010.00962.x

Benoit, L., Lehalle, H., & Jouen, F. (2004). Do young children acquire number words through subitizing or counting? *Cognitive Development*, *19*(3), 291-307. https://doi.org/10.1016/j.cogdev.2004.03.005

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology*, *42*(1), 189-201. https://doi.org/10.1037/0012-1649.41.6.189

Butterworth, B. (2010). Foundational numerical capacities and the origins of dyscalculia. *Trends in Cognitive Sciences*, *14*(12), 534-541. https://doi.org/10.1016/j.tics.2010.09.007

Bynner, J. M. (1997). Basic Skills in Adolescents' Occupational Preparation. *Career Development Quarterly*, *45*(4), 305-321. https://doi.org/10.1002/j.2161-0045.1997.tb00536.x

Cohen, D-J., & Blanc-Goldhammer, D. (2011). Numerical bias in bounded and unbounded number line tasks. *Psychonomic Bulletin & Review*, *18*(2), 331-338. https://doi.org/10.3758/s13423-011-0059-z

Desoete, A., Roeyers, H., & DeClercq, A. (2004). Children with mathematics learning disabilities in Belgium. *Journal of Learning Disabilities*, *37*, 50-6. https://doi.org/10.1177/00222194040370010601

Duchowski, A. T. (2017). *Eye Tracking Methodology: Theory and Practice*. Springer International Publishing. https://doi.org/10.1007/978-3-319-57883-5

Fischer, B., Gebhardt, C., & Hartnegg, K. (2008). Subitizing and Visual Counting in Children with Problems in Acquiring Basic Arithmetic Skills. *Optometry & Vision Development*, *39*(1), 24-29.

Garcia-Retamero, R., Andrade, A., Sharit, J., & Ruiz, J. G. (2015). Is patients' numeracy related to physical and mental health? *Medical Decision Making*, *35*(4), 501-511. https://doi.org/10.1177%2F0272989X15578126

Geary, D. C. (2011). Consequences, characteristics and causes of mathematical learning disabilities and persistent low achievement in mathematics. *Journal of Developmental and Behavioral Pediatrics*, *32*(3), 250-263. https://doi.org/10.1097/DBP.0b013e318209edef

Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., & Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development*, *78*(4), 1343-1359. https://doi.org/10.1111/j.1467-8624.2007.01069.x

Geary, D. C., Hoard, M. K., Nugent, L., & Byrd-Craven, J. (2008). Development of number line representations in children with mathematical learning disability. *Developmental Neuropsychology*, *33*(3), 277-299. https://doi.org/10.1080/87565640801982361

Gomez, A., Piazza, M., Jobert, A., Dehaene-Lambertz, G., & Huron, C. (2017). Numerical abilities of school-age children with developmental coordination disorder (DCD): A behavioral and eye-tracking study. *Human Movement Science*, *55*, 315-326. https://doi.org/10.1016/j.humov.2016.08.008

Gross-Tsur, V., Manor, O., & Shalev, R. S. (1996). Developmental dyscalculia: prevalence and demographic features. *Developmental Medicine and Child Neurology*, *38*(1), 25-33. https://doi.org/10.1111/j.1469-8749.1996.tb15029.x

Hamann, M. S., & Ashcraft, M. H. (1986). Textbook Presentations of the Basic Addition Facts. *Cognition & Instruction*, *3*(3), 173-202. https://doi.org/10.1207/s1532690xci0303_2

Heine, A., Thaler, V., Tamm, S., Hawelka, S., Schneider, M., Torbeyns, J., et al. (2009). What the eyes already 'know': using eye movement measurement to tap into children's implicit numerical magnitude representations. *Infant Child Development*., *19*(2), 175-186. https://doi.org/10.1002/icd.640

International Organization for Standardization. (2020). *Codes for the representation of names of countries and their subdivisions - Part 1: Country Code*. (ISO Standard No. 3166-1:2020).

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, *8*(4), 441-480. https://doi.org/10.1016/0010-0285(76)90015-3

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The discrimination of visual number. *The American Journal of Psychology*, *62*(4), 498-525. https://doi.org/10.2307/1418556

Koumoula, A., Tsironi, V., Stamouli, V., Bardani, I., Stavroula, S., Graham, A., Kafantaris, I., Charalambidou, I., Dellatolas, G., & von Aster, M. (2004). An Epidemiological Study of Number Processing and Mental Calculation in Greek Schoolchildren. *Journal of Learning Disabilities*, *37*(5), 377-388. https://doi.org/10.1177/00222194040370050201

Kucian, K., & von Aster, M. (2015). Developmental dyscalculia. *European Journal of Pediatrics*, *174*, 1-13. https://doi.org/10.1007/s00431-014-2455-7

Lai, M.-L., Tsai, M.-J., Yang, F.-Y., Hsu, C.-Y., Liu, T.-C., Lee, S. W.-Y., Lee, M.-H., Chiou, G.-L., Liang, J.-C., & Tsai, C.-C. (2013). A review of using eye-tracking

LUNDS UNIVERSITET

technology in exploring learning from 2000 to 2012. *Educational Research Review*, *10*, 90-115. https://doi.org/10.1016/j.edurev.2013.10.001

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

Lewis, C., Hitch, G. J., & Walker, P. (1994). The prevalence of specific arithmetic difficulties and specific reading difficulties in 9- to 10-year-old boys and girls. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *35*(2), 283-292. https://doi.org/10.1111/j.1469-7610.1994.tb01162.x

Lilienthal, A., & Schindler, M. (2019). Current trends in the use of eye tracking in mathematics education research: A PME survey. In M. Graven, H. Venkat, A. A. Essien, & P. Vale (Eds.), *Proceedings of 43rd Annual Meeting of the International Group for the Psychology of Mathematics Education*. Pretoria, South- Africa: PME.

Link, T., Nuerk, H. C., & Moeller, K. (2014). On the relation between the mental number line and arithmetic competencies. *The Quarterly Journal of Experimental Psychology*, *67*(8), 1597-1613. https://doi.org/10.1080/17470218.2014.892517

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

Mazzocco Michèle M. M., & Myers Gwen F. (2003). Complexities in Identifying and Defining Mathematics Learning Disability in the Primary School-Age Years. *Annals of Dyslexia*, *53*(1), 218-253. https://doi.org/10.1007/s11881-003-0011-7

Mejias, S., Grégoire, J., & Noël, M.-P. (2012). Numerical estimation in adults with and without developmental dyscalculia. *Learning and Individual Differences*, *22*, 164-170. https://doi.org/10.1016/j.lindif.2011.09.013

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

*Moeller, K., Neuburger, S., Kaufmann, L., Landerl, K., & Nuerk, H.-C. (2009). Basic number processing deficits in developmental dyscalculia: Evidence from eye tracking. *Cognitive Development*, *24*(4), 371-386. https://doi.org/10.1016/j.cogdev.2009.09.007

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Moser Opitz, E., Freesemann, O., Prediger, S., Grob, U., Matull, I., & Hußmann, S. (2016). Remediation for students with mathematics difficulties: An intervention study in middle schools. *Journal of Learning Disabilities*, *50*(6), 724-736. https://doi.org/10.1177/0022219416668323

Obersteiner, A., Reiss, K., Ufer, S., Luwel, K., & Verschaffel, L. (2014). Do first graders make efficient use of external number representations? The case of the twenty-frame. *Cognition and Instruction*, *32*(4), 353-373. https://doi.org/10.1080/07370008.2014.948681

Perttula, A. (2017). Eye tracking studies focusing on mathematics: a literature review. In *INTED2017. The 11th annual International Technology, Education and Development Conference: Valencia, 6th - 8th of March, 2017*. [2166-2173] IATED Academy. https://doi.org/10.21125/inted.2017.0639

Ramaa, S., & Gowramma, I. P. (2002). A systematic procedure for identifying and classifying children with dyscalculia among primary school children in India. *Dyslexia*, *8*(2), 67-85. https://doi.org/10.1002/dys.214

Rivera-Batiz, F. L. (1992). Quantitative Literacy and the Likelihood of Employment Among Young Adults in the United States. *Journal of Human Resources*, *27*(2), 313-328. https://doi.org/10.2307/145737

*Rottmann, T., & Schipper, W. (2002). Das Hunderter-Feld — Hilfe oder Hindernis beim Rechnen im Zahlenraum bis 100? *Journal für Mathematik-Didaktik*, *23*(1), 51-74. https://doi.org/10.1007/BF03338946

Salvucci, D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. Conference: *Proceedings of the Eye Tracking Research & Application Symposium*, *ETRA 2000*, Palm Beach Gardens, Florida, USA, November 6-8, 2000. https://doi.org/10.1145/355017.355028

*Schindler, M., Bader, E., Lilienthal, A. J., Schindler, F., & Schabmann, A. (2019a). Quantity Recognition in Structured Whole Number Representations of Students with Mathematical Difficulties: An Eye-Tracking Study. *Learning Disabilities: A*

Contemporary Journal, *17*(1), 5-28. http://urn.kb.se/resolve?urn=urn:nbn:se:oru:diva-79744

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

*Schindler, M., Schaffernicht, E., Lilienthal, A. (2019b) Differences in Quantity Recognition Between Students with and without Mathematical Difficulties Analyzed Through Eye: Analysis Through Eye-Tracking and AI In M. Graven, H. Venkat, A. Essien & P. Vale (ed.), *Proceedings of the 43rd Conference of the International Group for the Psychology of Mathematics Education* (pp. 281-288). PME.

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

Schleifer, P., & Landerl, K. (2011). Subitizing and counting in typical and atypical development. *Developmental Science*, *14*(2), 280-291. https://doi.org/10.1111/j.1467-7687.2010.00976.x

Schneider, M., Heine, A., Thaler, V., Torbeyns, J., De Smedt, B., Verschaffel, L., et al. (2008). A validation of eye movements as a measure of elementary school children's developing number sense. *Cognitive Development, 23*(3), 409-422. https://doi.org/10.1016/j.cogdev.2008.07.002

*Schot, W. D., van Viersen, S., van 't Noordende, J. E., Slot, E. M., & Kroesbergen, E. H. (2015). Strategiegebruik op de getallenlijntaak geanalyseerd met behulp van eye-tracking [Strategy use on the number line task analyzed using eye-tracking]. *Pedagogische Studiën*, *92*, 55-69.

Shalev, R. S., & Gross-Tsur, V. (2001). Developmental dyscalculia. *Pediatric Neurology*, *24*(5), 337-342. https://doi.org/10.1016/S0887-8994(00)00258-7

Shalev, R. S., Manor, O., & Gross-Tsur, V. (2005). Developmental dyscalculia: a prospective six-year follow-up. *Developmental Medicine and Child Neurology*, *47*(2), 121-125. https://doi.org/10.1017/s0012162205000216

Shalev, R. S., von Aster, M. (2008). Identification, classification, and prevalence of developmental dyscalculia. *Encyclopedia of Language and Literacy Development*: 1-9. https://doi.org/10.5167/uzh-12874

Siegler, R. S., & Booth, J. L. (2004). Development of Numerical Estimation in Young Children. *Child Development*, *75*(2), 428-444. https://doi.org/10.1111/j.1467-8624.2004.00684.x

Siegler, R., & Opfer, J. (2003). The development of numerical estimation: evidence for multiple representations of numerical quantity. *Psychological Science*, *14*, 237-243. https://doi.org/10.1111/1467-9280.02438

Skagerlund, K., Östergren, R., Västfjäll, D., & Träff, U. (2019). How does mathematics anxiety impair mathematical abilities? Investigating the link between math anxiety, working memory, and number processing. *PLoS ONE*, *14*(1), 1-17. https://doi.org/10.1371/journal.pone.0211283

Slusser, E. B., Santiago, R. T., & Barth, H. C. (2013). Developmental change in numerical estimation. *Journal of Experimental Psychology*, *142*(1), 193-208. https://doi.org/10.1037/a0028560

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

Strohmaier, A. R., MacKay, K. J., Obersteiner, A., & Reiss, K. M. (2020). Eye-tracking methodology in mathematics education research: A systematic literature review. *Educational Studies in Mathematics*, *104*, 147-200. https://doi.org/10.1007/s10649-020-09948-1

Sullivan, J. L., Juhasz, B. J., Slattery, T. J., & Barth, H. C. (2011). Adults' number-line estimation strategies: Evidence from eye movements. *Psychonomic Bulletin & Review*, *18*, 557-563. https://doi.org/10.3758/s13423-011-0081-1

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review*, *101*(1), 80-102. https://doi.org/10.1037/0033-295X.101.1.80

van der Sluis, S., de Jong, P. F., & van der Leij, A. (2004). Inhibition and shifting in children with learning deficits in arithmetic and reading. *Journal of Experimental Child Psychology*, *87*(3), 239-266. https://doi.org/10.1016/j.jecp.2003.12.002

*van der Weijden, F. A., Kamphorst, E., Willemsen, R. H., Kroesbergen, E. H., & van Hoogmoed, A. H. (2018). Strategy Use on Bounded and Unbounded Number Lines in Typically Developing Adults and Adults With Dyscalculia: An Eye-Tracking Study. *Journal of Numerical Cognition*, *4*(2), 337-359. https://doi.org/10.5964/jnc.v4i2.115

*van 't Noordende, J., van Hoogmoed, A., Schot, W., & Kroesbergen, E. (2016). Number line estimation strategies in children with mathematical learning difficulties measured by eye tracking. *Psychological Research*, *80*(3), 368-378. https://doi.org/10.1007/s00426-015-0736-z

*van Hoogmoed, A. H., van Wijk, D., van der Ven, S. H. G., & Kroesbergen, E. H. (n.d.). Number Magnitude Estimations of Children With and Without Mathematical Learning Disabilities on the Unbounded Number Line Task: An Eye Tracking Study [Unpublished manuscript]. Utrecht University.

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

van Wijk, D. R. (2017). *Unbounded Number Line Estimation Task: Quantity Estimations of Children With and Without Mathematical Learning Disabilities* (Master's thesis, Utrecht University, faculty of Social Sciences).

Wade, N. J. (2010). Pioneers of eye movement research. *i-Perception*, *1*, 33-68. https://doi.org/10.1068/i0389

Winoto, P., Tang, T. Y., Huang, Z., & Chen, P. (2017). “Thinking in Pictures?” Performance of Chinese Children with Autism on Math Learning Through Eye-Tracking Technology. In P. Zaphiris & A. Ioannou (Eds.), *Learning and Collaboration Technologies. Technology in Education* (pp. 215-226). Springer. https://doi.org/10.1007/978-3-319-58515-4_17

World Health Organization. (2019). *International statistical classification of diseases and related health problems* (11th ed.). https://icd.who.int/

LUNDS UNIVERSITET

Appendix

Table 2

Samples Included in the Systematic Review

#	Study	Sample	Comment
1.	• Schindler & Lilienthal (2018) • Schindler et al. (2019a) }	Based on the same participants and data.	Schindler & Lilienthal (2018) focused on eye tracking benefits over verbal reports. Schindler et al. (2019a) focused on differences in strategy-use.
2.	• Schot et al. (2015) • van Viersen et al. (2013) }	Schot et al. (2015) included one additional participant with mathematical difficulties. Otherwise the same participants.	Schot et al. (2015) used quantitative analysis of eye tracking data. Van Viersen et al. (2013) used qualitative analysis of eye tracking data.
3.	• van Hoogmoed et al. (n.d.)		
4.	• Moeller et al. (2009)		
5.	• van 't Noordende et al. (2016)		
6.	• Rottman & Schipper (2002)		
7.	• Schindler et al. (2020)		
8.	• van der Weijden et al. (2018)		
9	• Schindler et al. (2019b)		

Note. Table over the nine samples included in the systematic review. Square brackets are placed to illustrate which of the eleven studies that build upon the same, or almost the same empirical data.