EYE-TRACKING FOR STUDYING MATHEMATICAL DIFFICULTIES—ALSO IN INCLUSIVE SETTINGS

Maike Schindler[1] and Achim J. Lilienthal[2]

[1]University of Cologne and [2]Örebro University

Eye-Tracking (ET) is a promising tool for mathematics education research. Interest is fueled by recent theoretical and technical developments, and the potential to identify strategies students use in mathematical tasks. This makes ET interesting for studying students with mathematical difficulties (MD), also with a view on inclusive settings. We present a systematic analysis of the opportunities ET may hold for understanding strategies of students with MD. Based on an empirical study with 20 fifth graders (10 with MD), we illustrate that and why ET offers opportunities especially for students with MD and describe main advantages. We also identify limitations of think aloud protocols, using ET as validation method, and present characteristics of students' strategies in tasks on quantity recognition in structured whole number representations.

INRODUCTION

Eye-tracking (ET) promises to allow for online recording of cognitive processes (Chen & Yang, 2014). Its potential as a tool for mathematics education as well as important theoretical and technical developments have led to an increased interest in ET in the PME community, which the number of methodological papers and a strongly increased occurrence of ET-related key phrases in recent PME proceedings hint at.

The potential that ET may hold has been illustrated in many studies (e.g., Lindmeier & Heinze, 2016; Obersteiner & Tumpek, 2015). Some studies isolated differences between certain groups when solving the same math problems, e.g., differences between strong and weak students in mathematics (Rottmann & Schipper, 2002). Such research responds to the significance to understand mathematical difficulties. Not least since inclusive education has gained significance, research on students with mathematics difficulties (MD) has attracted more and more attention; with the aim to understand knowledge and learning in a fine-grained way and to foster students individually and adequately (e.g., Moser-Opitz et al., 2016; Scherer et al., 2016). Even though ET is partially already used with students with MD, research has—to the best of our knowledge—not yet systematically analyzed what opportunities ET holds especially for understanding strategies and thoughts of students with MD.

Thus, the aim of our study is to investigate opportunities ET may hold especially for students with MD. Based on an empirical study with 20 fifth graders (whereof 10 with MD), we investigated the potential benefit of ET in tasks addressing quantity recognition in structured whole number representations. We investigated the benefit of ET as compared to think aloud protocols (TA) and also compared the opportunities of ET for students with and without MD. The results of our data analysis indicate that ET

2018. In NNN (Eds.). *Proceedings of the 42nd Conference of the International Group for the Psychology of Mathematics Education*, Vol. 1, pp. XX-YY. Umeå, Sweden: PME.

is especially advantageous for students with MD, whose strategies appear to be more diverse and, for several reasons, more difficult to explain. Our paper contributes to mathematics education research in three major ways: We illustrate that and why ET offers opportunities especially for students with MD and describe its main advantages. We identify limitations of TA for students with MD using ET as validation method. And, finally, we present strategies of students with and without MD in tasks on quantity recognition in structured whole number representations.

THEORETICAL BACKGROUND

Eye-Tracking (ET) and its use in mathematics education research

ET devices aim to identify gaze points by capturing eye movements and projecting from the fovea onto the surrounding scene. Video-based systems currently dominate the market; either in the form of head-mounted devices (ET glasses) or remote devices attached to a computer screen that displays visual stimuli (Holmqvist et al., 2011). ET devices promise to allow for online recording of cognitive processes (Chen & Yang, 2014). However, ET only provides a flickering view on "shadows" cast by brain processes in the form of eye movements. Accordingly, interpretation of ET data is non-trivial. It typically rests on the "eye-mind" hypothesis, which, as expressed by (Just & Carpenter, 1976) in the reductionist spirit of a brain-computer metaphor, posits that "the eye fixates the referent of the symbol currently being processed if the referent is in view. That is, the fixation may reflect what is at the 'top of the stack'" (p. 441). Data interpretation is challenging since the eye-mind hypothesis does not always hold (Holmqvist et al., 2011). Not all cognitive processes are tightly linked to visual stimuli. Also, foveal vision is not always required, e.g. when peripheral vision is sufficient.

Despite the challenges mentioned, ET is a potent tool for mathematics education research (MER). Powerful inferences are possible especially in specific, controlled settings—e.g., in "visually presented cognitive tasks" (Obersteiner & Tumpek, 2015, p. 257), and by using domain-specific interpretation—i.e. considering known semantics of fixated visual entities (Schindler & Lilienthal, 2017). The improved theoretical and computational means for interpretation as well as the advent of commercial, less intrusive, portable and increasingly affordable ET devices (Holmqvist et al., 2011) led to considerable interest in this technology; not least in the PME community. When analyzing the last five PME conferences, we found 25 ET-related papers, mostly ET studies and a smaller but increasing number of methodological papers dedicated to ET and the interpretation of eye movement data. We also carried out a full text analysis of the proceedings of the last five PME conferences and found a clear trend: While the wordings "eye tracking" and "eye movement" occurred only 20 times at PME37, this number increased to 208 at PME41.

In addition to improving ET devices and better means of interpretation, interest in ET in MER is fueled by its potential "as a method for identifying strategy use in mathematical tasks" (Beitlich & Obersteiner, 2015). This paper contributes to identify opportunities (and limitations) of ET for identifying strategy use in mathematical tasks. This is particularly important in comparison to methodological alternatives for the same purpose, including TA methods and response time tests.

Mathematics difficulties (MD) and their identification

Learning difficulties in mathematics are an important topic in practice and research and have attracted increased interest not least since inclusive education has gained significance. However, to date there is no consensus on a definition or term characterizing the group of students having difficulties in mathematics (Scherer et al., 2016). Certain researchers speak of *mathematical learning disabilities*, others of (severe) *mathematical difficulties*, depending on different national educational contexts and research traditions (see ibid.; Moser-Opitz et al., 2016). In this paper, we follow Moser-Opitz et al. (2016) who talk about students with mathematics difficulties (MD) and summarize their potential difficulties, comprising (see also Scherer et al., 2016): verbal counting (e.g., counting by groups and counting principles), grouping, degrouping, the base-10 number system and understanding the place value, understanding the meaning of operations, solving word problems; as well as factual knowledge, fact retrieval and (deficits in) working memory.

To support students with MD in their mathematical learning, researchers and teachers aim to identify students' individual assets, difficulties, and strategies. Mathematics education, special education, and psychology use different methods for diagnosing: They consider, e.g., written products or tests, observations, or processing times of students when working on math problems. These methods have demonstrated their potential but are also limited because they consider an "outside" view on mental processes or use end results, which can make it difficult to distinguish between different processes leading to the same products. TA and ET, however, observe manifestations of internal processes and have great potential to identify strategy use.

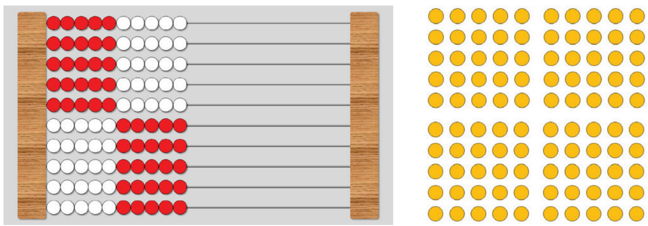
Despite its advantages and popularity, TA has drawbacks, especially when working with children with MD and other learning difficulties. In Concurrent TA (CTA) where participants are asked to verbalize their thoughts while performing a task, the additional cognitive load can be overwhelming (Ericsson & Simon, 1980), especially for students with MD for whom the task itself and verbalization can constitute a major cognitive effort. Instead, in Retrospective TA (RTA) verbalization occurs after completion of the task. Still, aspects that may affect verbalizations of students with MD in RTA may be, e.g., anxiety, difficulties with memory retrieval, introspection, or meta-cognitive reflection, or verbalization issues.

We see that ET has two potential major advantages over TA: It can identify unconscious processes and there is no verbalization step that could absorb cognitive resources and potentially influence students' strategies. To the best of our knowledge, the potential of ET for better understanding strategies of students with MD has not yet been empirically investigated.

We ask the following research questions: *What opportunities may ET offer for understanding strategies of students with MD?* We approach this question through an empirical study with students in an inclusive education setting. We use ET—in combination with TA and without—also address the questions *What opportunities does ET offer compared to TA?* and *May the benefit of ET for understanding students' strategies be bigger for students with MD than for students without MD?*

METHOD

*Students. F*or answering the research questions, we use data from a research project with 20 fifth-grade students, ages between 9;11 and 11;11, in a German comprehensive school ("Gesamtschule"). The participating school was in a town of 80,000 inhabitants, situated on the edge of a German urban area. The study took place in the first weeks of fifth grade, after the students had finished the German primary school after grade 4. Students with mathematical difficulties (MD) were identified through qualitative diagnostic interviews addressing MD (following Schulz & Wartha, 2012) investigating, e.g., students' number sense and understanding of number and operations (see Moser-Opitz et al, 2016). Among the 10 students with MD, there were four with special educational needs (in learning, social and emotional development, and physical development).



100-bead abacus 100-dot square

Figure 1: Number representations used in this study

Tasks. The students worked on tasks to determine the number of beads/dots on a 100-bead abacus and on a 100-dot square (Fig. 1). We first let the students determine the total number of beads in a 100-bead abacus (number of dots in the 100-dot square) and asked them to think aloud after they had given their answers (immediately after each task). Further, the students determined numbers (e.g., 7, 76, and 92) on the 100-bead abacus (the 100-dot square) without TA. In this set of tasks, we wanted to analyze students' strategies without potential interference of TA. Although structured number representations on the 100-dot square and abacus are addressed in second grade in Germany, we used this kind of tasks in grade 5, following Moser-Opitz et al. (2016) who point out, "some research (…) shows that low achievers in mathematics in the higher grades lack very basic competencies, such as counting in groups or understanding the base-10 system, even with small numbers" (p. 1f.). In previous studies, ET has proven to be useful to analyze students' strategies working with such structured representations. This is due to the fact that the representations comprise visually presented information and ET can help to understand how students capture such information. Lindmeier and Heinze (2016) analyzed students' strategies to determine numbers of dots/beads and found significant differences between first graders and adults, in particular different strategy use (e.g., counting, subitizing, or using structures). They conclude that their "study shows that eye-tracking data can be used to access different strategies when solving number tasks in structured representations" (p. 7). Rottmann and Schipper (2002) compared high and low achievers' use of the 100-dot square in addition and subtraction tasks using ET.

Analyzing scanpaths, they found, e.g.: “Low achievers (…) use (…) [the material] in a way which turns out not to be helpful at all for them: their activities are either inappropriate or are unilaterally subordinated to the counting-all strategy” (p. 51).

Eye-Tracker. We used the ET glasses Tobii Pro Glasses 2 with a framerate of 50 Hz. They weigh 45 grams, are relatively unobtrusive, allow for reliable tracking, and are robust to students touching them. Even though tasks were presented on a computer screen, we decided not to use a remote eye-tracker to be able to capture students’ eye movements when looking away from the screen, and also gestures, e.g. pointing.

Data analysis. We focus on the gaze-overlaid videos (videos taped by the ET glasses including the eye gazes as dot wandering around in the video; similar to Schindler & Lilienthal, 2017). The data analysis followed Mayring’s (2014) qualitative content analysis in an inductive manner. The first three steps were (1) *description* of the student’s *eye movements* in the video, (2) *paraphrasing* the content-bearing semantic elements in the description relevant for identifying student strategies *and transposing* them to a uniform stylistic level, and (3) *category development*, i.e. inductively assigning categories to the data with according descriptions/definitions. After having categorized all data, we went through the data once more in a *category revision* step; which included partially re-categorization. In a final subsumption step, we collected all instances matching every category. We also compared the informative content of ET and TA. For this purpose, we analyzed the TA for every student. We transcribed utterances and gestures (e.g., pointing) and then used the same steps as outlined above.

RESULTS

Comparing ET and TA

In the tasks to determine the number of beads (dots) on a 100-bead abacus (100-dot square), the students were asked to think aloud how they found out the number. When comparing eye movements and verbal descriptions of the students, three scenarios appeared (Tab. 1). 11 students’ data (6 with MD, 5 without) were analyzed (the others were excluded, e.g., due to data loss or interference of the interviewer). Most cases matched scenario A, where ET provided more information about students’ strategies.

	ET more informative	ET/TA equally informative	TA more informative
Scenario	A	B	C
Number of cases	6 (5)	4 (1)	1 (0)

Table 1: Scenarios of informative content of ET vs. TA (numbers of students with MD in brackets)

Scenario A. Johanna, a student with MD and special needs in social and emotional development, in particular with anxieties, participated in this study (she herself was eager to take part). When determining the number of dots in the 100-dot square, her eyes glissaded over the first 5 dots in the first row and then over the second five in the first row, indicating that she grasped the number of dots in the first row. Then, she looked at the left edge of every row one after the other (beginning at the top) indicating that she was counting rows. She then said “100”. When being asked “How did you do this?” by the interviewer, she answered “I don’t know”. When asked once more, she answered “5, 10, 15”, which is hardly conclusive. In this case, ET appears to be more informative than TA for understanding the student’s strategy. Johanna might have been anxious to explain her strategy (e.g., in order to avoid failure) or not used to explain her thoughts. In other cases, we experienced that the students had difficulties to express their thoughts (e.g., due to poor language skills or shyness). In yet another case, Jasmine did not mention a strategy that, as the ET revealed, she had used (counting 5s). When being asked about it, she said: “No, I first counted the 25s”. Maybe she forgot about the strategy she used initially; or intentionally denied the strategy—possibly because she perceived that counting is not an expected strategy. ET as compared to TA was not only beneficial when students did not answer or denied a strategy. It furthermore often provided a greater level of detail and reflected processes that the TA only hinted at; for instance how (where and in which order) the students really counted or how they perceived ones. Besides, TA required the interviewer to ask follow-up questions which may guide students’ strategies and implicitly transport certain norms with an impact on the students’ behavior.

Scenario B. In four cases, ET and TA were equally informative. In Samuel’s case, ET indicated that he first grasped the number of beads in the first row of the abacus (saccade over the 10 beads) and then counted the number of rows (subsequent fixations on the middles of the rows one after the other). Verbally, he described: “It is 10 beads per rod. And then I simply, like 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 (pointing to the rows one after the other in the middles)”. His description contained the same information as the ET. Daniel and Elena also used pointing and speech explaining their strategies. It appears that pointing has the potential to compensate for flaws in verbalization (e.g., about the order of counted elements, the exact focus, etc.).

Scenario C. In one case, Simon’s (a student without MD), eye movements were very brief (several saccades over the dot field) and thus partially ambiguous for the data analysis. In TA, he was able to express himself well and used additional pointing. In this case, TA and ET did not contradict each other, but—given the briefness of Simon’s eye movements—TA provided less ambiguous information on his strategy.

Using ET videos only

Students’ use of structures and strategy use. ET gave indications on whether students used, e.g., the structure of 10, 5, or 50. We found that less than half of the students in the MD group but nearly all in the control group used the 50-structure in the abacus (for quantities such as 54 or 68). Besides, in the MD group, there were three students that only used the structures of 5 (Nidal and Jasmine) or the structure of 10 (Ava) in the 100-dot square. Except for a few instances where ET data were ambiguous, we could assign strategies to eye movements. Strategies included, e.g., counting in structures (e.g. rows) and then determining the ones through, e.g., subitizing; or subtraction strategies (e.g. for 92: recognizing the 8 remaining beads and inferring 92). We found that 14 out of 20 students used (at least once) more than one strategy working on single problems, whereas the students in TA always only explained one single strategy (possibly due to certain norms in the math classroom).

Explanation of processing times. The ET data analysis gave hints why students' processing times (partially measured as reaction times) may be prolonged. Reasons for longer processing times include: use of different strategies (e.g., hedging, or realizing that another strategy is advantageous), repeating the same strategies (e.g., re-counting, sometimes in different orders), use of time-consuming strategies (e.g., counting 5s or counting ones one by one), or slow execution of strategies (e.g., when counting). Besides, we noticed students looking around on non-meaningful entities, which may be caused, e.g., by stress (e.g., because students realize an issue).

Explanation of student mistakes. The students made several mistakes when determining the numbers of beads and dots. Analyzing ET videos helped us to understand why students made such mistakes and that the same wrong answers may have different reasons. For instance, the result "99" instead of 89 may appear because students wrongly grasp the number of rows, or because they make an error in a subtraction strategy. Even though ET cannot always entirely clarify students' inferences, ET appears to be a helpful tool to understand where mistakes originate.

DISCUSSION

The power of ET was already shown in several studies in MER. However, the potential it holds especially for students with MD was not systematically analyzed yet. For this reason, we conducted an empirical study with 20 students (10 with MD) with the aim to investigate the opportunities ET may hold especially for students with MD.

We investigated the *benefit of ET as compared to TA* and found that in most cases, ET provided more detailed information, and appeared to be especially beneficial for analyzing strategies of students with MD. TA was in some cases equally informative as ET. In these cases, gestures and pointing typically helped the students to express their strategies. This hints at the significance of students' ability to express themselves for valid TA; which appears to be a generally important factor for analyzing strategies of students, e.g. with a different mother tongue and migration background. Our results suggest that students should be trained and explicitly asked to use pointing in TA to increase the informative content, when ET cannot be used. In the tasks without TA, ET revealed that many students used more than one strategy to solve the tasks, which they never reported in the TA condition. This indicates that students' verbal reports in TA may not reflect the variety of strategies students would have used without TA.

We also compared *opportunities of ET for students with and without MD*. We found that for most students with MD (5 out of 6), ET provided more (detailed) information than TA, whereas this was only the case for a minority of the students without MD. Possible reasons are that students with MD used multiple and more diverse strategies more often, which increases the difficulty to verbalize. We assume that explanation difficulty but also anxieties (due to disadvantageous prior experiences) and normative aspects (e.g., hesitance to explain counting strategies because it may not be appreciated) may have an influence in this respect.

However, our study is only a small scale study with only one kind of tasks. Its results cannot and should not easily be generalized. However, it hints at an important fact: That ET may be especially valuable for students with MD, where TA is particularly difficult. We believe our research to be a springboard for further research on this topic.

References

Beitlich, J. T. & Obersteiner, A. (2015): Eye-Tracking as a Method for Identifying Mathematical Strategies. In Beswick, K., Muir, T., & Fielding-Wells, J. (Eds.) (2015). *Proceedings of the 39th Conference of the International Group for the Psychology of Mathematics Education, Vol. 1,* 89-93. Hobart, Australia: PME.

Chen, Y.C. & Yang, FY. (2014) Probing the Relationship Between Process of Spatial Problems Solving and Science Learning: An Eye Tracking Approach. *International Journal of Science and Mathematics Education*, *12*(3), 579-603.

Ericsson, A. & Simon, H. A. (1980). Protocol Analysis: Verbal Aspects as Data. *Psychological Review, 87*(3), 215-251.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). Eye *tracking: A comprehensive guide to methods and measures*. Oxford: OUP.

Just, M.A. & Carpenter, P.A. (1976). Eye Fixations and Cognitive Processes. *Cognitive Psychology*, 8, 441-480.

Lindmeier, A., & Heinze, A. (2016). *Strategies for recognizing quantities in structured whole number representations – A comparative eye-tracking study.* Paper presented at 13th International Congress on Mathematical Education (ICME-13), 2016.

Mayring, P. (2014). Qualitative content analysis: theoretical foundation, basic procedures and software solution. Klagenfurt, Austria: Beltz.

Moser Opitz, E., Freesemann, O., Prediger, S., Grob, U., Matull, I., & Hußmann, S. (2017). Remediation for students with mathematics difficulties: An intervention study in middle schools. *Journal of learning disabilities*, *50*(6), 724-736.

Obersteiner, A. & Tumpek, C. (2015). Measuring fraction comparison strategies with eye-tracking. *ZDM—Mathematics Education*, 48(3), 255-266.

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.

Scherer, P., Beswick, K., DeBlois, L., Healy, L., & Opitz, E. M. (2016). Assistance of students with mathematical learning difficulties: how can research support practice?. *ZDM*, *48*(5), 633-649.

Schindler, M. & Lilienthal, A. (2017). Eye-Tracking and its Domain-Specific Interpretation. A Stimulated Recall Study On Eye Movements In Geometrical Tasks. In Proceedings of PME 41, Vol. 4, 153-160.

Wartha, S., & Schulz, A. (2012). Rechenproblemen vorbeugen. Grundvorstellungen aufbauen. Zahlen und Rechnen bis 100. Berlin: Cornelsen.