



Investigating middle school students' eye movements on the mathematical representations: An eye-tracking study

Hilmi Karaca¹ · Erhan Ertekin² · Kursat Cagiltay³

Received: 26 July 2024 / Accepted: 10 February 2025 / Published online: 26 February 2025
© The Author(s) 2025

Abstract

In mathematics education, representations are used in place of mathematical structures, ideas, or relationships to concretize, transform, and represent them. When students interact with these representations, they engage in various cognitive activities such as thinking, reasoning, understanding, remembering, problem-solving, attention, and decision-making, which are difficult to observe. Therefore, uncovering these cognitive activities is very significant for mathematics education. However, they are not easy to uncover as they cannot be directly observed. Eye tracking is an important approach that can be used to reveal cognitive activities that cannot be directly observed. This study investigated how middle school students examine representations by examining their eye movements. Eighty-five (40 girls and 45 boys) 7th-grade middle school students participated in the study. In the study, gaze durations, fixation count, and fixation duration on four different representation types: verbal representation, symbolic representation, number line representation, and counters representation were compared. The findings showed that students fixated more on the verbal representation and gazed at it for longer. However, fixation durations on the verbal representation were quite short compared to the other representations. In contrast, when examining the counters, there were fewer fixations and shorter gaze durations, but fixation durations were longer. Gazes on the number line and symbolic representation did not differ across all three variables. The findings indicated that gaze on verbal and non-verbal representations differed to some extent, but not entirely. Finally, the findings are discussed in the context of mathematical representation and eye-tracking literature.

Keywords Representation · Mathematical representation · Eye-tracking · Integer · Mathematics education

This study is derived from the corresponding author's doctoral dissertation.

Extended author information available on the last page of the article

1 Introduction

Many representations and representation systems have been used in the development process of mathematics throughout history. While some of these are still in use, others have been abandoned over time. Numbers, symbols, graphs, diagrams, models, verbal expressions, geometric figures, etc. are some of the mathematical representations used today and are the result of many years of cultural refinement. The word meaning of the concept of representation is a specially created structure that corresponds to all or part of something (Palmer, 1977), replacing ideas, meanings, or relationships (Gilbert, 2010). It is used in making sense of, using, and transferring any information (Goldin, 2014). However, a representation may not mean the same thing for every individual, or a representation may be used in different meanings. For this reason, representations carry meaning within personal or culturally structured systems (Goldin & Kaput, 1996). This is because internally, there is an act of interaction and interpretation in the relationship between the representer and the represented (von Glasersfeld, 1987). Moreover, Duval (1999) stated that no representation of a mathematical object can fully reflect the mathematical object and that the representation can only be used to talk about the mathematical object.

The use of representation in mathematics education is based on the understanding and use of mathematical ideas (NCTM, 2000). Representations in mathematics teaching are structures that replace, embody, encode, and represent mathematical ideas or relationships in the most common sense (Goldin, 2014). For this reason, the use of representations is extremely important in mathematics education to organize and convey mathematical ideas, select and use appropriate mathematical representations, and model and interpret problem situations (NCTM, 2000).

The fact that the concept of representation is handled from different aspects has led to the classification of representations in various ways. In the most general sense, Goldin and Kaput (1996) divided the representations that can be directly observed with sensory organs such as symbols, formulas, and concrete manipulatives into external/physical representations and the structures and associations that an individual constructs in his/her mind while solving a problem into internal/mental representation. Lesh et al. (1987) defined five different representational systems in the context of mathematics learning and problem-solving. These are real-world situations, manipulative models, pictures, spoken language, and written symbols. When this classification of representations is examined, it is seen that Lesh et al. (1987) classify representations within the framework of observable structures. Gilbert (2010) analyzed representations as verbal and non-verbal representations within the framework of Paivio's (1990) dual coding theory. Accordingly, voice, speech, written expressions, and symbols were categorized as verbal representations, while figures, graphics, and pictures were categorized as non-verbal representations. While verbal representations are relational, hierarchical and generally have clearer meaning, non-verbal representations are more open to interpretation and have a part-whole relationship (Gilbert, 2010; Paivio, 1990).

1.1 Eye tracking in mathematics education

The examination of eye movements can provide a perspective on the cognitive processes of the individual (Just & Carpenter, 1976). Eye trackers record pupil movements and identify the points of gaze (Carter & Luke, 2020). Thus, it helps to collect information about a person's visual attention (Duchowski, 2017; Lilienthal & Schindler, 2019). Eye-tracking is a research method that has attracted increasing interest in mathematics education research in terms of the opportunities it offers (Lai et al., 2013; Lilienthal & Schindler, 2019; Schindler & Lilienthal, 2019; Strohmaier et al., 2020). This is because, apart from traditional research methods, the eye-tracking method offers the opportunity to examine research topics in detail and understand cognitive processes (Lilienthal & Schindler, 2019). When the studies conducted within the scope of mathematics education are examined, it is observed that there is a significant increase in the number of eye-tracking studies (Strohmaier et al., 2020).

It was observed that the previous studies were generally related to numbers and geometry. Counting, single-digit and multi-digit numbers, magnitude-smallness in numbers, spatial-numerical relationship, number representations, basic arithmetic operations, rational numbers and ratio and proportion were the prominent topics (Atagi et al., 2016; Bolden et al., 2015; Susac et al., 2014), while geometric proof, spatial movements, geometric construction were the prominent topics in geometry (Alqassab et al., 2018; Lee & Wu, 2018; Lin & Lin, 2014; Schindler & Lilienthal, 2019). In addition, there are a limited number of eye-tracking studies on reasoning and proof, the use of mathematical representation, learning difficulties, computer-assisted learning, mathematical problem-solving, and statistics (Lin, 2024; Strohmaier et al., 2020).

Where we gaze and how long we gaze are influenced by cognitive processes such as attention, perception, memory, decision-making, etc. (Just & Carpenter, 1980). Just and Carpenter (1980) explained this interaction with the eye-mind hypothesis (EMH). According to this hypothesis, as long as a stimulus (a word, a letter, an image, or a part of an image) is processed in the mind, the eye is fixed on that stimulus. In other words, “visual focus equals cognitive focus” (Strohmaier et al., 2020). However, Schindler and Lilienthal (2019) stated that although visual attention and cognitive attention overlap in general, EMH may not always be valid. For this reason, although the EMH provides a framework for determining mental functioning, it is important to investigate the mental process to be investigated within a context and to analyze and interpret the data of eye movements within this context (Schindler & Lilienthal, 2019).

Representations provide important advantages in mathematics education in terms of supporting different ideas and thoughts and providing in-depth understanding (Ainsworth et al., 1997). When the mathematical representations used in teaching integers and operations with integers are analyzed, number line representation and counters representation come to the fore (Bozkurt & Polat, 2011; Koç-Şanlı & Işık, 2020). This is because students perceive different types of representations differently. In some of the studies involving symbolic and visual representations, it was observed that students focused longer on symbolic representations (Beitlich et al., 2014; Ott et al., 2018), while in others, students focused longer on visual representations (Beitlich

et al., 2015; Malone et al., 2020). Although there are different reasons for this situation, criteria such as the nature of the participants, the structure and comprehensibility of the symbols, and the simplicity or complexity of the visuals, may have caused the differentiation of focusing between symbols and visuals. These results show that there is a difference between the perception of verbal and non-verbal representations within the scope of the representation classification made by Paivio (1990) and Gilbert (2010).

In addition to the opportunities it offers in mathematics education, eye tracking is a common approach used to obtain information about cognitive processes and to measure and analyze cognitive load (Alemdag & Cagiltay, 2018; Cong et al., 2022; Joseph & Muruges, 2020). Cognitive load theory refers to the amount of mental effort or resources required to complete a given task (Chandler & Sweller, 1991; Sweller, 1988), based on the assumption that working memory capacity is limited (Huh et al., 2019). In this context, cognitive load theory provides a perspective on the use of cognitive resources in cognitive activities such as attention, learning and problem solving (Chandler & Sweller, 1991). The learner's expertise and the nature of the material being learned (intrinsic cognitive load), the design of the learning materials and the learning environment (extraneous cognitive load), or the construction of new schemata and their association with prior knowledge (germane cognitive load) can all create cognitive load in the learner's mind. Intrinsic and extraneous cognitive load tend to negatively affect learning, whereas germane cognitive load tends to positively affect learning, and the sum of these cognitive loads constitutes total cognitive load and is limited to working memory (Plass et al., 2010). Eye tracking metrics such as gaze duration, fixation duration, fixation number, saccade length, blink duration and pupil dilation are some of the eye tracking data used to measure cognitive load (Chien et al., 2015; Cong et al., 2022; Susac et al., 2017).

1.2 Aim of the study

The previous related studies are generally based on student responses and subjective data obtained from interviews. In addition, while comparatively examining how students use different types of representations, we have not come across a study that investigates how they analyze these representations in terms of representation. Since the eye-tracking method provides important data that cannot be directly observed, it provides very useful and important data in the analysis of mathematical representations. In this context, this study aims to investigate how middle school students' eye movements on the representations used in addition with integers differ according to representation types. For this purpose, counters representation (CR), number line representation (NLR), symbolic representation (SR), and verbal representation (VR), which are the four representations commonly used in addition with integers, were used in the study. Eye movements on the representations were analyzed both within each representation and by making comparisons between the representations. During the examination of the literature, it was observed that various types of representations were commonly compared under different conditions (Andrà et al., 2009; Andrà et al., 2015; Atagi et al., 2016; Beitlich et al., 2014; Bolden et al., 2015; Dewolf et al., 2015; Lee & Wu, 2018; Liang & She, 2021; Malone et al., 2020; Norqvist et al.,

2019; Ott et al., 2018; Susac et al., 2017). However, none of these studies specifically examined CR, NLR, or SR in isolation. Therefore, examining eye movements within the representation distinguishes this study from other studies. The findings obtained were used to answer the following questions.

- Is there a difference between the participants' eye movement measurements on the CR?
- Is there a difference between the participants' eye movement measurements on the NLR?
- Is there a difference between the participants' eye movement measurements on the SR?
- Do participants' eye movement measurements differ according to representation type?

2 Method

2.1 Participants

As operations with integers constitute a sub-area in the seventh-grade mathematics curriculum defined by the Turkish Ministry of National Education (TMNE) (2018), our study focused on seventh-grade middle school students. We randomly selected eighty-five students, with forty females and forty-five males from three different middle schools in the Central Anatolia region of Türkiye, all of whom were seventh-grade students. To ensure more accurate results from the eye tracker, participants were specifically chosen among students who did not wear glasses.

In eye-tracking studies, the quality of eye movement data can significantly impact the research (Duchowski, 2017). In this study, we developed an “Eye Tracking Data Checklist” to assess the quality of the collected data. The checklist was constructed using data extraction criteria found in previous research (Alqassab et al., 2018; Atagi et al., 2016; Beitlich et al., 2014; Bulf et al., 2014; Miller-Singley & Bunge, 2018; Obersteiner & Tumpek, 2016; Sullivan et al., 2011). After assessing the data collected from eighty-five participants using the Eye Tracking Data Checklist, it was identified that the data of twenty-one participants were deemed unsuitable and were consequently excluded from the study before the analysis. The analysis was carried out on the data of sixty-four participants, consisting of thirty girls and thirty-four boys.

2.2 Materials

In eye-tracking applications, we modelled integer addition using SR, VR, CR, and NLR. We selected these representations because they are the most used formats in integer operations and were designed in alignment with Paivio's (1990) dual coding theory of representations. The dual coding theory comprises two representation systems: the verbal representation system and the non-verbal representation system (Gilbert, 2010; Paivio, 1990). According to this classification, the VR and SR used in the

study fall under the verbal representation system, while the CR and NLR are categorized under the non-verbal representation system. Furthermore, four area of interests (AOIs) were generated on CR (CRa, CRb, CRc and CRd) and SR (SRa, SRb, SRc and SRd) and three AOIs were generated on NLR (NLR0, NLRa and NLRb), but no AOI was generated on VR (Fig. 1). In CRs and SRs, the position of the AOIs do not change, whereas in NLRs, the zero point is labelled NLR0, the tip of the first arrow is labelled NLRa, and the tip of the second arrow is labelled NLRb.

To enable more in-depth investigations of the representations, we restricted the operations to addition. We generated a total of twelve questions, with three questions stemming from each representation. Various pairs of integers were employed in addition operations, with attention to using one-digit addition questions to enhance mental calculation ease. Additionally, a broad range of addition operations was incorporated, encompassing different number pairs (positive+positive, positive+negative, negative+positive, negative+negative) whenever possible. In the questions, the solutions to the addition operations were provided, while the results of the operations were omitted to encourage participants to concentrate more on the representations. Students were expected to provide answers as part of the task.

Different design features of stimuli such as color, size, and location can affect eye movements (Duchowski, 2017; Livingston et al., 1992). The representations were designed, considering the principles in the theory of multimedia learning (ML) as presented by Mayer (2009). These principles, individual differences, and cognitive load are important factors affecting eye movement measurements (Coskun & Cagiltay, 2022). While designing the representations, we considered the principles of consistency and redundancy. Furthermore, we incorporated the principle of personalization in VR. In this context, we ensured that the representations used in the presentations had a consistent color scheme (white background with black representations), similar sizes, an absence of other stimuli alongside the representations, and

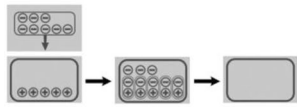
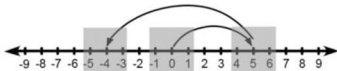
Operation	Representation	AOIs
CR [(+5)+(-8)]		CRa CRb CRc CRd
NLR [(+5)+(-9)]		NLRb NLR0 NLRa
SR [(-8)+(+7)]	$(-8) + (+7) = [(-1) + (-7)] + (+7)$ $= (-1) + [(-7) + (+7)]$ $= ?$	SRa SRb SRc SRd
VR [(+7)+(-3)]	Pozitif yedi ile negatif üçün toplanması işleminde; sayılar ters işaretli olduğu için verilen sayıların mutlak değer karşılıkları olan yediden üç çıkarılır. Mutlak değeri daha büyük olan sayının işareti göz önüne alındığında, işlemin sonucu olur.	

Fig. 1 Examples of representation and AOIs

that similar types of representations were uniform (for example, the arrows on NLR were of the same type, and the same font was used in VR). We utilized the Arial font for both letters and numbers.

2.3 Apparatus, environment, and calibration

In the study, the GazePoint GP3 eye-tracking device was used to determine participants' eye movements. The GP3 is a research-grade screen-based eye-tracking device and uses a machine vision camera with a 60 Hz sampling rate at the center of the imaging and processing system. It also has 0.5–1.0 degree visual angle accuracy, 60 Hz sampling rate, 25 cm × 11 cm (horizontal × vertical) mobility, and 5–9 point calibration support. The eye-tracking device was used in combination with a monitor with a screen width of 17 inches and a screen resolution of 1024 × 768. The calibration process and presentation of stimuli were conducted using the Gazepoint Analysis Professional Edition software.

The data collection was conducted in the schools where the participants were studying. Locations such as libraries or computer laboratories with minimal background noise were selected to minimize distractions for the participants. The lighting in the environment was kept as low as possible to optimize the recording of eye movements. The 9-point eye calibration feature provided by Gazepoint Analysis Professional Edition was used to determine participants' gaze points. The calibration process typically lasted about 1 min for each participant.

2.4 Procedures

Before commencing the study, all participants received a comprehensive introduction to the research. Those who expressed their desire to take part in the study were provided with a voluntary participation form and parental consent was sought when applicable. Participants who completed the required consent documents were eligible to participate. Before the actual eye-tracking procedure, participants were briefed on the procedure through an informational document, and the eye-tracking device was introduced. To acquaint participants with the study, a trial run was conducted on paper using similar questions that were not included in the actual application.

A pilot study was conducted to determine how long the stimuli would appear on the screen to prevent unnecessary eye movements. The pilot study found that the majority of participants answered the questions within 30 s. Therefore, an upper limit of 30 s was set for each stimulus. The presentations were set to automatically switch to the next one after 30 s. However, if participants answered the questions in a shorter time, the presentation of the stimuli was adjusted so that they could move on to the next question by pressing the “space” key if they wished. Furthermore, to eliminate the order effect, the presentation of the twelve representations was presented in a shuffled order to each participant during the eye-tracking procedure. Stefkovics and Kmetty (2022), refer to the order effect as a phenomenon in which the experiences of previous questions affect the cognitive response process and responses to subsequent questions. The order effect is a factor that can affect the quality of eye movement data due to more attentive attention to the first questions, distraction towards the end,

fatigue or desensitization during the process (Duchowski, 2017; Höhne & Lenzner, 2015; Kohout et al., 2023). To avoid the order effect, it was decided to present the questions in a shuffled order during the actual eye-tracking procedure, and the questions were presented to each of the students in a different order. In this context, care was taken to ensure that the same representation types did not come one after the other, that a representation type was not concentrated at the beginning or at the end, and that the representation types were homogeneously distributed.

Participants were given two tasks to complete before the eye-tracking procedure for the questions presented to them. These tasks included: “In each presentation, examine the representations and determine the result of the operation” and “Verbally express the result of the operation, then proceed to the next presentation by pressing the space bar on the keyboard.” Participants were instructed to minimize any extraneous movements, such as getting closer to or farther from the screen or turning their heads left or right, and to maintain their gaze on the screen as much as possible. During the eye-tracking procedure, no interventions were made, and participants who requested it were provided with the opportunity to review their eye movement recordings at the end of the procedure.

2.5 Data analysis

To compare participants' eye movements, we computed the average gaze duration (AGD), average fixation count (AFC), and average fixation duration (AFD) for each participant's eye movements on each representation. The averages of the measurements correspond to the average of the measurements obtained from the same representations. These data were calculated both for the representation as a whole and the AOI created on each representation separately. When comparing between representations, we used data for the representation as a whole, while when analyzing within representations, we used data for the AOI.

In the analyses, the independent variables include the AOI within each representation for within-representation comparisons and the four different representation types (CR, NLR, SR, and VR) for between-representation comparisons. The dependent variables for the analyses consist of measures obtained from the AOI and the representation as a whole. To assess the normal distribution of the data, skewness, and kurtosis values were examined, and data were considered to be normally distributed when the skewness and kurtosis coefficients were within the $(-2, +2)$ range (George & Mallery, 2016).

Since the data were collected from a single group and there were more than two data groups, one-factor ANOVA for repeated measures (RM ANOVA) was used to compare parametric data. RM ANOVA is a quantitative data analysis method employed to ascertain if there is a significant difference among the means of more than two data groups acquired from the same sample (Pallant, 2020). When a difference in means was identified through ANOVA, we employed the Bonferroni multiple comparison test to identify which data points contributed to the difference. In the comparison of nonparametric data, the Friedman test, which is the nonparametric equivalent of one-factor ANOVA, was used and the Wilcoxon test was used to determine which measurements the difference was between.

3 Results

The graphs of eye movements on the representations are given in Fig. 2. AGD and AFC were observed more in VR and less in CR. On the contrary, AFD was found less in VR and more in CR. NLR and SR data were remarkably close to each other.

The AFC data on CRa, CRb, and SRc, the AFC data on CR, and the AFC data on CRc, NLRb, SRa, SRb, and NLR are not normally distributed (Table 1). For this reason, RM ANOVA was employed to compare AFC measurements within CR AOIs, AGD and AFC measurements within NRL AOIs, AFC measurements within SR AOIs, and AGD measurements across representation types. The Friedman test was used to compare AGD and AFD measurements within CR AOIs, AFD measurements within NRL AOIs, AGD and AFD measurements within SR AOIs, and AFC and AFD measurements across representation types. Furthermore, when the assumption of sphericity was not met in the RM ANOVA data, epsilon corrections were implemented. Specifically, the Huynh-Feldt correction was used for $\epsilon > 0.75$, and the Greenhouse-Geisser correction was applied for $\epsilon < 0.75$ (Can, 2013).

The results of the Friedman test reveal a significant difference among the AGD values related to CR's AOIs ($\chi^2_{(2, n=64)}=74,625, p<0,05$) (Table 2). The Wilcoxon signed-rank test indicates statistically significant differences between all AGD values (CRa-CRb, CRa-CRc, and CRb-CRc) ($p<0.05$) (Table 3). Similarly, there is a significant difference in the AFD values ($\chi^2_{(2, n=64)}=64,406, p<0,05$), with the Wilcoxon signed-rank test showing that the differences between all AFD values (CRa-CRb, CRa-CRc, and CRb-CRc) are statistically significant ($p<0.05$).

RM ANOVA results suggest a significant difference in AFC values ($F_{(1,273-80,176)}=86,297, p<0,05$) (Table 4). According to the Bonferroni multiple comparison test, there is a significant difference between CRa-CRc and CRb-CRc ($p<0.05$) (Table 5). In summary, participants spent more time gazing at CRc compared to CRa, while AFC and AFD values were higher in CRc compared to the others. However, AFC and AFD values in CRa and CRb are quite similar to each other. It's important to note that 59% of the participants did not exhibit any eye movements in the CRd AOIs, which is why this area was not included in the analysis.

RM ANOVA results indicate a significant difference in AFC values for NLR's AOIs ($F_{(1,55-99,638)}=61,99, p<0,05$). Similarly, a significant difference was observed in the AFC values ($F_{(1,64-103,312)}=34,462, p<0,05$). According to the Bonferroni mul-

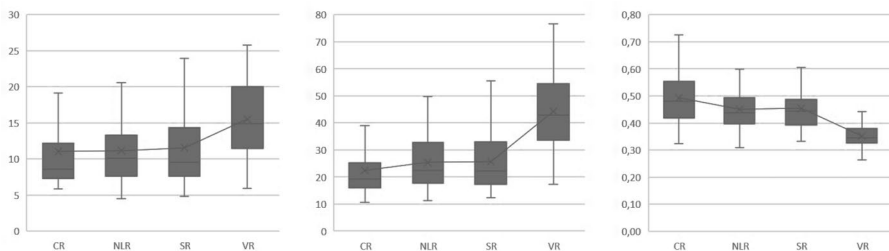


Fig. 2 Graphs of eye movement data by representation types

Table 1 Descriptive statistics of data for measurements on representations and AOIs

Reps.*	Measurements	AOIs	N	M (S.D.)	Median**	Skewness (S.E.)	Kurtosis (S.E.)
CR	AGD (sec)	CRa	64	1,89 (1,27)	1,63	1,37 (0,3)	2,12 (0,59)
		CRb	64	2,29 (1,78)	1,93	1,56 (0,3)	2,43 (0,59)
		CRc	64	5,83 (3,02)	4,8	0,94 (0,3)	0,16 (0,59)
	AFC (n)	CRa	64	5,31 (2,89)		1,14 (0,3)	1,030 (0,59)
		CRb	64	5,56 (3,05)		1,26 (0,3)	1,626 (0,59)
		CRc	64	10,54 (4,99)		1,15 (0,3)	1,421 (0,59)
	AFD (sec)	CRa	64	0,33 (0,65)	0,33	0,67 (0,3)	0,45 (0,59)
		CRb	64	0,38 (0,88)	0,34	0,66 (0,3)	-0,07 (0,59)
		CRc	64	0,57 (1,69)	0,54	3,04 (0,3)	16,51 (0,59)
NLR	AGD (sec)	NLR0	64	1,74 (0,93)		0,87 (0,3)	-0,01 (0,59)
		NLRa	64	3,88 (2,25)		1,14 (0,3)	1,46 (0,59)
		NLRb	64	3,26 (1,23)		0,5 (0,3)	0,54 (0,59)
	AFC (n)	NLR0	64	6,67 (2,74)		0,79 (0,3)	0,5 (0,59)
		NLRa	64	9,92 (5,3)		1,23 (0,3)	1,59 (0,59)
		NLRb	64	7,65 (2,96)		0,87 (0,3)	0,64 (0,59)
	AFD (sec)	NLR0	64	0,26 (0,064)	0,25	0,59 (0,3)	0,95 (0,59)
		NLRa	64	0,38 (0,07)	0,37	0,73 (0,3)	0,22 (0,59)
		NLRb	64	0,47 (0,189)	0,41	2,33 (0,3)	8,67 (0,59)
SR	AGD (sec)	SRa	64	2,95 (2,1)	2,42	0,9 (0,3)	0,33 (0,59)
		SRb	64	3,09 (2,17)	2,34	1,11 (0,3)	0,83 (0,59)
		SRc	64	3,19 (2,82)	2,3	1,64 (0,3)	4,03 (0,59)
	AFC (n)	SRa	64	9,89 (5,95)		0,63 (0,3)	-0,38 (0,59)
		SRb	64	10,4 (6,05)		0,89 (0,3)	0,39 (0,59)
		SRc	64	9,19 (5,82)		0,71 (0,3)	0,2 (0,59)
	AFD (sec)	SRa	64	0,27 (0,11)	0,24	1,27 (0,3)	2,61 (0,59)
		SRb	64	0,27 (0,09)	0,28	0,89 (0,3)	3,45 (0,59)
		SRc	64	0,3 (0,12)	0,27	0,81 (0,3)	0,5 (0,59)
All Reps.	AGD (sec)	CR	64	11,03 (5,37)		1,43 (0,3)	1,08(0,59)
		NLR	64	11,17 (4,58)		1,03 (0,3)	0,56 (0,59)
		SR	64	11,59 (5,61)		1,35 (0,3)	1,16 (0,59)
		VR	64	15,52 (5,16)		0,23 (0,3)	-0,91 (0,59)
	AFC (n)	CR	64	22,42 (9,55)	19,33	1,58 (0,3)	2,43 (0,59)
		NLR	64	25,39 (10,25)	22,5	1,25 (0,3)	1,31 (0,59)
		SR	64	25,66 (1075)	22,17	1 (0,3)	0,14 (0,59)
		VR	64	44,32 (14,17)	42,83	0,07 (0,3)	-0,6 (0,59)
	AFD (sec)	CR	64	0,49 (0,1)	0,48	0,55 (0,3)	-0,2 (0,59)
		NLR	64	0,45 (0,08)	0,44	1,52 (0,3)	4,16 (0,59)
		SR	64	0,45 (0,08)	0,45	1,72 (0,3)	5,6 (0,59)
		VR	64	0,35 (0,04)	0,35	0,6 (0,3)	0,21 (0,59)

*Representations, **Median value is given for non-normally distributed data

tuple comparison test, statistically significant differences ($p < 0.05$) exist between AGD and AFC measurements for all AOIs (NLRa-NLRb, NLRa-NLR0, NLRb-NLR0). Friedman test results show differences in AFD measures related to NLR's AOIs ($\chi^2_{(2, n=64)} = 91,219, p < 0,05$). The Wilcoxon signed-rank test results indicate

Table 2 Friedman test results

Representations	Measurements	<i>N</i>	Chi-Square	df	<i>p</i>
CR	AGD	64	74,625	2	0,00*
	AFD	64	64,406	2	0,00*
NLR	AFD	64	91,219	2	0,00*
SR	AGD	64	4,29	2	0,117
	AFD	64	0,976	2	0,614
Between Representations	AFC	64	100,413	3	0,00*
	AFD	64	124,735	3	0,00*

p*<0,05Table 3** Wilcoxon comparison test results

Representations	Measurements	Compared AOs / Representations	<i>Z</i>	<i>p</i>
CR	AGD	CRa-CRb	2,083	0,037*
		CRa-CRc	6,708	0,00*
		CRb-CRc	6,681	0,00*
	AFD	CRa-CRb	2,869	0,004*
		CRa-CRc	6,647	0,00*
		CRb-CRc	5,47	0,00*
NLR	AFD	NLRa-NLR0	-	0,00*
			6,955	
		NLRb-NLR0	-	0,00*
	AFC		6,888	
		NLRb-NLRa	-	0,001*
			3,394	
Between Representations	AFC	CR-NLR	2,598	0,009*
		CR-SR	2,729	0,006*
		CR-VR	6,902	0,00*
		NLR-SR	0,524	0,6
		NLR-VR	6,868	0,00*
	AFD	SR-VR	6,838	0,00*
		CR-NLR	3,842	0,00*
		CR-SR	4,548	0,00*
		CR-VR	6,955	0,00*
		NLR-SR	0,383	0,701
		NLR-VR	6,867	0,00*
		SR-VR	6,942	0,00*

p*<0,05Table 4** RM ANOVA results

Representations	Measurements	<i>N</i>	df (conditions/error)	<i>F</i>	<i>p</i>	η_p^2
CR	AFC	64	1,273/80,176**	86,297**	0,00*	0,578
NLR	AGD	64	1,55/99,638***	61,99***	0,00*	0,496
	AFC	64	1,64/103,312***	34,462***	0,00*	0,354
SR	AFC	64	1,657/104,397***	948***	0,376	0,015
Between Representations	AGD	64	3/189	25,444	0,00*	0,288

p*<0,05, **Greenhouse-Geisser, * Huynh-Feldt

Table 5 Bonferroni comparison test results

Representations	Measurements	Compared AOI / Representations	Difference between means (S.E.)	<i>p</i>
CR	AFC	CRa-CRb	-0,250 (0,24)	0,906
		CRa-CRc	-5,234 (0,56)	0,00*
		CRb-CRc	-4,984 (0,48)	0,00*
NLR	AGD	NLRa-NLRb	0,618 (0,24)	0,033*
		NLRa-NLR0	2,136 (0,21)	0,00*
		NLRb-NLR0	1,518 (0,14)	0,00*
NLR	AFC	NLRa-NLRb	2,271 (0,45)	0,00*
		NLRa-NLR0	3,255 (0,45)	0,00*
		NLRb-NLR0	0,984 (0,29)	0,003*
Between Representations	AGD	CR-NLR	-0,145 (0,6)	1,00
		CR-SR	-0,558 (0,55)	1,00
		CR-VR	-4,490 (0,66)	0,00*
		NLR-SR	-0,413 (0,63)	1,00
		NLR-VR	-4,345 (0,57)	0,00*
		SR-VR	-3,932 (0,59)	0,00*

* $p < 0,05$

statistically significant differences between all AFD values (NLRa-NLRb, NLRa-NLR0, NLRb-NLR0) ($p < 0.05$).

The results of the comparisons of measurements on SR according to AOIs indicate that there is no significant difference between the AGD ($\chi^2_{(2, n=64)}=4,29$), AFC ($F_{(1,657-104,397)}=948$), and AFD ($\chi^2_{(2, n=64)}=0,976$) values ($p > 0.05$). Therefore, AGD, AFC, and AFD measurements on SR do not vary significantly across AOIs. It's important to note that 43% of the participants did not exhibit any eye movements in the SRd, and therefore, this AOI was not included in the analysis.

RM ANOVA results reveal a significant difference between the AGD values across all representations ($F_{(3-189)}=25,444$, $p < 0,05$). According to the Bonferroni multiple comparison test, there is a statistically significant difference between the AFC values of VR and the other representations ($p < 0.05$). Similarly, the Friedman test results show a significant difference in AFC ($\chi^2_{(3, n=64)}=100,413$) and AFC ($\chi^2_{(3, n=64)}=124,735$) values across all representations ($p < 0.05$). Wilcoxon signed-rank test results indicate a significant difference in pairwise comparisons, except for the NRL-SR pairwise comparisons ($p > 0.05$), for both AFC and AFD values.

4 Discussion and Conclusion

In this section, we first discuss the findings of each representation according to their AOIs, followed by the results of comparisons between the representations.

4.1 Discussion and conclusion for the CR

The study found that participants' focus was concentrated on CRc, largely due to its role in the addition process and the higher number of counters present. This concentration of eye movements on CRc resulted in increased cognitive load (Chandler & Sweller, 1991; Sweller, 1988), as reflected in the higher AGD and AFD measures observed in this area. In the context of cognitive load, longer AFD measures indicate deeper cognitive processing and greater task difficulty (Just & Carpenter, 1980). High AFD measures are also considered to be indicative of significant cognitive effort required for information processing (Hahn & Klein, 2022). Similarly, the AGD can be used as a measure of cognitive load. For example, Susac et al. (2017) highlighted that graphs and tables help to organize data, reducing the time students spend on irrelevant details and redirecting their focus to more important information. This process leads to a reduction in AGD, minimizing unnecessary use of cognitive resources and ultimately contributing to a reduction in cognitive load during data interpretation. The increased AGD and AFD values observed in CRc suggest that this area required sustained attention and higher order processing, consistent with the notion that cognitive load increases in areas requiring greater mental effort (Hahn & Klein, 2022; Just & Carpenter, 1980; Susac et al., 2017).

In comparing CRa and CRb, it was found that while AFC values were similar, AGD and AFD were longer in CRb. This was contrary to expectations, as CRa had more counters and was presumed to require more time. The findings suggest that participants engaged in more complex cognitive processes while focusing on CRb, indicating that factors beyond the mere number of counters influence eye movement and cognitive load.

The results of the study also provide insight into the order of attentional shifts between different AOIs. The highest and lowest AGD and AFD measures occurred in CRc-CRb-CRa, respectively, and statistically significant differences were observed between these measures. When analyzing the gaze paths of the participants' eye movements, three different basic gaze patterns come to the fore. The first is the CRa-CRb-CRc gaze sequence. Here, parallel to the reading direction (Carter & Luke, 2020; Rayner, 2009), participants performed addition as a cognitive activity by focusing on CRc after looking at CRa and CRb, which contain the numbers to be added, starting from the upper left corner. The second is the CRa-CRb gaze. Here, the participant performed the addition while looking at CRb after receiving the first number to be added from CRa. While cognitive activities took place while looking at CRb, looking at CRc was more limited here. Third, the gaze was focused directly on CRc. Here, the participant performed most of the cognitive activities only in CRc, while looking at other AOIs for a very short time. This indicates that they recognized that the information in CRc was sufficient for the addition process and aimed to directly reach the result of the operation without wasting time on other AOIs. The AFD and AGD measures of the AOIs are also consistent with these rankings. It can be said that CRc is critical for CR and that this AOI is usually the end point of the gazes. However, participants who did not want to focus on the number and density of the stamps in CRc were able to perform addition using only the data in CRa and CRb. It can be said that these shifts of attention caused more cognitive activity to

occur in CRb compared to CRa and led to an increase in AGD and AFD measures in CRb. Furthermore, CRa-CRb-CRc and CRa-CRb gaze shifts may increase the amount of information stored in memory over time, leading to increased AGD and AFD measures due to limited working memory capacity (Huh et al., 2019; Plass et al., 2010), thus increasing cognitive load in CRb and CRc. Overall, eye movements in CR were more intense in CRc compared to CRa and CRb, indicating the importance of CRc for cognitive activities in CR. In addition, it also seems that cognitive load may increase in CRb with the information obtained from CRa.

4.2 Discussion and conclusion for the NLR

Analysis of eye movements in the AOIs generated in the NLR revealed significant differences due to both information integration and familiarity. Participants exhibited significantly less attention and focus at NLR0, as evidenced by lower measures of AGD, AFC and AFD compared to other AOIs. The low level of focus in this AOI could be attributed to the inherent informational sparsity of NLR0 or participants' long-term familiarity with its content, which they acquired during primary education, possibly reducing cognitive demand (Chandler & Sweller, 1991; Chien et al., 2015; Cong et al., 2022). Because NLR0 has a lower density of information compared to other AOIs because it represents a starting point on the number line, and students encounter similar representations from the early years of primary school education. These findings are consistent with Sweller's (1988) claim that existing knowledge structures facilitate information processing and thus reduce cognitive load.

In contrast, NLRa emerged as an important area for knowledge integration, with high AGD and AFC values. The convergence of arrows representing numerical inputs in NLRa requires information synthesis by engaging working memory subsystems responsible for temporal storage and the visuospatial dimension (Baddeley, 2003a, b). Repetitive fixations (high AFC) and prolonged gaze (high AGD) signal iterative cognitive efforts to associate numerical data derived from the visual, consistent with integrative processing demands in visuospatial tasks (Baddeley, 2003a, b; Holmqvist et al., 2011).

However, NLRb elicited significantly longer AFD, indicative of intensive cognitive processing. This prolonged fixation duration is consistent with System 2 cognitive processes involving deliberate analytical reasoning, such as mental arithmetic or result verification, as described by Evans (2003) and Kahneman (2011). System 2 is associated with measures of general intelligence that are relatively slower and deliberate, allowing for abstract and hypothetical thinking, limited by working memory capacity (Evans, 2003; Kahneman, 2011). Moreover, NLRb, positioned as the final node in the NLR, can be said to further increase cognitive process intensity by requiring participants to allocate additional attentional resources for error checking or final decision making (Just & Carpenter, 1980). These findings partially contradict Bolden et al. (2015). This difference observed from differences in gaze durations (here the NLRa was gazed at longer) may be due to methodological differences such as task complexity or participant demographics.

In summary, in NLRa, visual spatial construction (intersection of number line arrows) may require frequent reorganization and checking to distinguish between

nested and implicit data. In NLRb, we think that the fact that it is last in the sequential structure increases fixation durations due to the aim of reducing the risk of error and the presence of the result arrow. In addition, tasks that emphasized accuracy rather than speed may have led participants to consciously focus more on the NLRb. These inferences underline the interaction between the structure of the NLR and participant strategies and attentional allocations. In conclusion, the NLR findings demonstrate how AOI-specific cognitive demands ranging from information integration (NLRa) to analytic calculation (NLRb) are reflected in eye movement metrics. Furthermore, the roles of both intrinsic (e.g., information density and information integration) and extrinsic (e.g., task design) factors in the context of cognitive processes are examined by relating visual-numerical processing to eye-tracking evidence.

4.3 Discussion and conclusion for the SR

No statistically significant differences were found between the AGD, AFC, and AFD measures of AOIs in the SR. In contrast to dynamic or complex representations, the tightly organized structure of the SR is thought to reduce the need for integration across AOIs by standardizing cognitive processes such as comprehension and addition. This structural clarity is likely to enable automatic processing (Sweller, 1988), suggesting that participants make sense of the representation with similar eye movements across AOIs of the SR and generate solutions by relying on their existing schemas. This automaticity is further supported by the lack of significant differences in AGDs, AFCs, and AFDs across the AOIs. In addition, SR, which is characterized by its symbolic structure, typically provides clear, simple, understandable, and limited-interpretation information (Gilbert, 2010; Paivio, 1990). These inferences suggest that SR reflects gaze and attentional strategies that are evenly distributed across interests, rather than gaze focused on a particular AOI, and that AOIs generate similar cognitive load (Baddeley, 2003a; Chandler & Sweller, 1991).

On the other hand, it can be argued that eye movement measures in SR AOIs indicate optimized resource allocation rather than uniform cognitive passivity. This is because, unlike other representations, each AOI in the SR (except SRd) had sufficient independent information to perform the given task. That is, the AOIs of the SR were functionally equivalent to each other. This design of the SR provides a structure in which familiar symbols can be processed as a single unit in each AOI, providing units that require equivalent information processing and interpretation activities. In the context of eye-tracking measures, it can be argued that participants categorized and filtered information with a similar structure (Wickens, 2021) in this process to reduce working memory load (Chandler & Sweller, 1991; Huh et al., 2019).

However, the observed homogeneity in eye-tracking measures may also reflect limitations in the design of the SR or task parameters. The SR structure and task instructions may have predisposed participants to superficial processing, thereby masking implicit cognitive strategies. Beyond automaticity and facility, future studies could investigate cognitive activities on the SR by differentially organizing the SR's structure and task instructions and eliminating identities in the SR's AOIs.

4.4 Discussion and conclusion for representation types

The analysis of eye movements across different representation types revealed significant insights into how students perceive and process mathematical representations. Notably, the AGD results indicated that participants spent more time fixating on VR than on other representations. This suggests that the cognitive activities associated with similar operations are more time-consuming when using VR. This finding aligns with certain studies in literature, while contradicting others. For instance, Atagi et al. (2016) observed shorter AGD for visual representations compared to SR. Ott et al. (2018) conducted a study involving multiple representations and found that more AFC and longer AGD were observed in the text representation compared to the formula and graph representations. Conversely, Andr   et al. (2015) reported shorter AGD on formulas compared to graphs, whereas Beitlich et al. (2015) found that more time was spent on visual representations than on symbolic and text representations. The results from this study corroborate the findings of Atagi et al. (2016) and Ott et al. (2018), while contrasting with those of Andr   et al. (2015) and Beitlich et al. (2015). However, no significant differences were observed in AGD across the other representations. This suggests that cognitive processes such as reading, comprehending, analyzing, and decoding require similar amounts of time for comparable addition operations in SR, CR, and NLR. The increased AGD for VR implies that more transformations are involved when transitioning from external representations to internal representations in VR (Goldin & Kaput, 1996). This is due to the interaction and interpretation that occur between the representer and the represented, both internally and externally (von Glasersfeld, 1987). Furthermore, although VR contains more comprehensible words compared to CR, NLR, and SR, the sequential, relational, and hierarchical structure of these words (Paivio, 1990) increases the gaze time, potentially prolonging the time required to reach the same cognitive stage using VR.

The highest AFC was recorded for VR, while the lowest AFC was observed for CR, similar to the findings of Ott et al. (2018). Ott et al. (2018) concluded in their study with multiple representations that more AFC and AGD occurred in the text representation compared to the formula and graphical representations. The increased engagement with the text representation was attributed to students positioning it as a reference representation. In this study, the longer duration of gazing at VR and the higher AFC may be attributed to the large number of pieces of information (e.g., each word or phrase) present in VR. However, this result contradicts the findings of Malone et al. (2020).

While the least AFD occurred on VR, the highest AFD was observed on CR. Longer fixation durations generally indicate a more comprehensive cognitive process (Rayner, 1998). However, longer fixations can also occur when participants encounter meaningless words or complex problems (Hegarty et al., 1992, 1995; Liang & She, 2021). Accordingly, the longer AFD on CR compared to other representations suggests that participants engaged in more intensive cognitive processing on CR. Furthermore, according to the Eye-Mind Hypothesis, where and how long we gaze is influenced by cognitive processes (Just & Carpenter, 1980; Schindler & Lilienthal, 2019), and these activities generate cognitive load in the mind (Chandler & Sweller, 1991; Huh et al., 2019; Plass et al., 2010; Sweller, 1988). This may indicate that visual representations like CR increase cognitive load, similar to findings by Andr   et al. (2015) and Ott et al. (2018).

Studies involving verbal and non-verbal representations have reported varying results regarding fixation durations. Some found longer fixations on VR (Beitlich et al., 2014; Liang & She, 2021; Ott et al., 2018), while others observed longer fixations on non-verbal representations (Beitlich et al., 2015; Malone et al., 2020). In this study, fixations were longer on CR (non-verbal representation) and shorter on VR. Additionally, when comparing NLR and SR, no statistically significant differences were found in AGD, AFC, and AFD between these representations. This suggests a partial differentiation in AGD for verbal and non-verbal representations, influenced by factors such as participant characteristics, text length and complexity, and visual simplicity or complexity.

In summary, this study reveals that AGD on verbal representations (SR, VR) and non-verbal representations (CR, NLR) generally differ. AGD increases while AFD decreases in VR and SR, whereas in visual representations like NLR and CR, AGD is shorter while AFD is longer. These findings indicate that students' perception, understanding, and interpretation of representations are influenced by the types of representations used. Consequently, the differentiation in eye movements between verbal and non-verbal representations is partial but not complete, highlighting the complexity of cognitive processes involved in interpreting different mathematical representations.

5 Recommendations and limitations

The conclusions demonstrate that students perceive and analyze various types of mathematical representations differently. Moreover, each representation type confers distinct advantages contingent upon its contextual application and situational relevance. Based on these conclusions, the following recommendations are proposed to enhance pedagogical strategies and optimize learning outcomes:

- Teachers should carefully consider segments within mathematical representations that require heightened cognitive load, such as CRc within CR and NLRb within NLR. Providing structured learning activities that gradually introduce, and scaffold complex tasks can help students manage cognitive load effectively.
- Researchers could explore the effectiveness of metacognitive strategies in guiding students' attention and cognitive processing across different AOIs within the number line representation. Because investigating how metacognitive awareness impacts eye movement patterns and fixation durations in NLR can provide valuable insights into optimizing instructional practices.
- Teachers should adopt differentiated instructional strategies that acknowledge the distinct cognitive demands of verbal (SR, VR) and non-verbal (CR, NLR) representations. For instance, tasks involving VR could focus on enhancing visual literacy and interpretation skills through interactive simulations or multimedia resources. Similarly, activities centered on CR and NLR should emphasize hands-on manipulatives or digital tools that facilitate spatial reasoning and numerical understanding. By tailoring teaching approaches to each representation type's cognitive characteristics, teachers can foster deeper engagement and comprehension among students.

- Researchers could conduct studies exploring how individual differences, such as affective features or prior knowledge, influence eye movements and cognitive processing across different representation types. Understanding these differences can provide insights into personalized learning approaches to teaching these representations that meet different student needs.
- Given the varying patterns observed in AGD, AFC, and AFD across representation types, teachers should integrate multimedia learning environments that combine verbal and non-verbal elements. Such environments can leverage the strengths of each representation type while promoting holistic cognitive engagement. This approach not only strengthens conceptual understanding but also prepares students to adapt their problem-solving strategies across different representation types.
- This study provides some key findings for future research on the subject. In particular, it presents unique results for CR, NLR and SR, and it has been observed that these representations have not been sufficiently analyzed in previous studies. For this reason, there is a need for studies in which these representations are analyzed in more detail.

This study provides some basic findings for future research on the subject. It shows that more research should be conducted by diversifying the types of representations, participant groups, and the subject of the study. In particular, conducting more comprehensive studies in larger samples by enriching them with different methods may increase the reliability of the results.

As a result, it shows that there is a significant relationship between students' perceptions of representations, but the available data and the results obtained are not sufficient to understand and explain the reasons for this relationship. In addition, the results are limited by the number of participants, the types of representations, and the subject matter used in the representations. Therefore, caution should be taken about the generalizability of the results. More in-depth and comprehensive research needs to be conducted to understand the reasons for the relationships and to reach more general conclusions.

Acknowledgements None.

Author contributions HK: Conceptualization, Methodology, Investigation, Data curation, Statistical analysis, Data interpretation, Writing—original draft, Writing—review and editing. EE: Conceptualization, Statistical analysis, Data interpretation, Supervision, Writing—review and editing. KC: Conceptualization, Data interpretation, Supervision, Writing—review and editing.

Funding Open access funding provided by the Scientific and Technological Research Council of Türkiye (TÜBİTAK).

No financial support was received from any institution or organization for the study.

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

Declarations

Ethics approval and consent to participate All ethical standards were considered and complied with during the research. The compliance of the study with ethical principles was approved by the Necmet-

tin Erbakan University Social Sciences and Humanities Scientific Research Ethics Committee and the approval number is 2021/10–541. Participation in the study was carried out on a completely voluntary basis. Participants who agreed to participate in the study and completed the voluntary participation form were included in the study.

Consent for publication We consent the publication of this manuscript in Education and Information Technologies.

Competing interests The research was conducted without any commercial or financial relationship that could be interpreted as a potential conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ainsworth, S. E., Bibby, P. A., & Wood, D. J. (1997). Information technology and multiple representations: New opportunities - new problems. *Journal of Information Technology for Teacher Education*, 6(1), 93–105. <https://doi.org/10.1080/14759399700200006>
- Alemdag, E., & Cagiltay, K. (2018). A systematic review of eye tracking research on multimedia learning. *Computers & Education*, 125, 413–428. <https://doi.org/10.1016/j.compedu.2018.06.023>
- Alqassab, M., Strijbos, J. W., & Ufer, S. (2018). The impact of peer solution quality on peer-feedback provision on geometry proofs: Evidence from eye-movement analysis. *Learning and Instruction*, 58, 182–192. <https://doi.org/10.1016/J.LEARNINSTRUC.2018.07.003>
- Andrá, C., Lindström, P., Arzarello, F., Holmqvist, K., Robutti, O., & Sabena, C. (2015). Reading mathematics representations: An eye-tracking study. *International Journal of Science and Mathematics Education*, 13(S2), 237–259. <https://doi.org/10.1007/s10763-013-9484-y>
- Andrá, C., Arzarello, F., Ferrara, F., Holmqvist, K., Lindstrom, P., Robutti, O., & Sabena, C. (2009). How students read mathematical representations: An eye tracking study. In M. Kaldrimidou & H. Sakonidis (Eds.), *Proceedings of the 33rd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 49–56). Thessaloniki, Greece: PME.
- Atagi, N., DeWolf, M., Stigler, J. W., & Johnson, S. P. (2016). The role of visual representations in college students' understanding of mathematical notation. *Journal of Experimental Psychology: Applied*, 22(3), 295–304. <https://doi.org/10.1037/xap0000090>
- Baddeley, A. (2003a). Working memory and language: An overview. *Journal of Communication Disorders*, 36(3), 189–208. [https://doi.org/10.1016/S0021-9924\(03\)00019-4](https://doi.org/10.1016/S0021-9924(03)00019-4)
- Baddeley, A. (2003b). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4(10), 829–839. <https://doi.org/10.1038/nrn1201>
- Beitlich, J. T., Obersteiner, A., Moll, G., Mora Ruano, J. G., Pan, J., Reinhold, S., & Reiss, K. (2014). The role of pictures in reading mathematical proofs: An eye movement study. In P. Liljedahl, C. Nicol, S. Oesterle, & D. Allan (Eds.), *Proceedings of the 38th Conference of the International Group for the Psychology of Mathematics Education and the 36th Conference of the North American Chapter of the Psychology of Mathematics Education* (Vol. 2, pp. 121–128). Vancouver, Canada: PME.
- Beitlich, J. T., Obersteiner, A., & Reiss, K. (2015). How do secondary school students make use of different representation formats in heuristic worked examples? An analysis of eye movements. In K. Beswick, T. Muir, & J. Wells (Eds.), *Proceedings of the 39th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 97–104). Hobart, Australia: PME.

- Bolden, D., Barnby, P., Raine, S., & Gardner, M. (2015). How young children view mathematical representations: A study using eye-tracking technology. *Educational Research*, 57(1), 59–79. <https://doi.org/10.1080/00131881.2014.983718>
- Bozkurt, A., & Polat, M. (2011). Teachers' views on the effect of modeling with counters on learning integers. *Gaziantep Üniversitesi Sosyal Bilimler Dergisi*, 10(2), 787–801.
- Bulf, H., Cassia, V. M., & de Hevia, M. D. (2014). Are numbers, size and brightness equally efficient in orienting visual attention? Evidence from an eye-tracking study. *Plos One*, 9(6), 1–6. <https://doi.org/10.1371/JOURNAL.PONE.0099499>
- Carter, B. T., & Luke, S. G. (2020). Best practices in eye tracking research. *International Journal of Psychophysiology*, 155, 49–62. <https://doi.org/10.1016/j.ijpsycho.2020.05.010>
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4), 293–332. https://doi.org/10.1207/s1532690xc0804_2
- Chien, K. P., Tsai, C. Y., Chen, H. L., Chang, W. H., & Chen, S. (2015). Learning differences and eye fixation patterns in virtual and physical science laboratories. *Computers & Education*, 82, 191–201. <http://doi.org/10.1016/j.compedu.2014.11.023>
- Cong, R., Tago, K., & Jin, Q. (2022). Measurement and verification of cognitive load in multimedia presentation using an eye tracker. *Multimedia Tools and Applications*, 81(19), 26821–26835. <https://doi.org/10.1007/s11042-022-13294-0>
- Coskun, A., & Cagiltay, K. (2022). A systematic review of eye-tracking-based research on animated multimedia learning. *Journal of Computer Assisted Learning*, 38(2), 581–598. <https://doi.org/10.1111/jcal.12629>
- Dewolf, T., Van Dooren, W., Hermens, F., & Verschaffel, L. (2015). Do students attend to representational illustrations of non-standard mathematical word problems, and, if so, how helpful are they? *Instructional Science*, 43(1), 147–171. <https://doi.org/10.1007/s11251-014-9332-7>
- Duchowski, A. T. (2017). Eye tracking methodology. Springer International Publishing. <https://doi.org/10.1007/978-3-319-57883-5>
- Duval, R. (1999). Representation, vision and visualization: Cognitive functions in mathematical thinking. Basic issues for learning. In F. Hitt, & M. Santos (Eds.), *Proceedings of the 21st PME*, 3–26. Mexico: Cuernavaca, Morelos.
- Evans, J., & St., B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, 7(10), 454–459. <https://doi.org/10.1016/j.tics.2003.08.012>
- George, D., & Mallery, P. (2016). *IBM SPSS Statistics 23 Step by Step: A Simple Guide and Reference* (14th ed.). Routledge. <https://doi.org/10.4324/9781315545899>
- Gilbert, J. K. (2010). The role of visual representations in the learning and teaching of science: An introduction. *Asia-Pacific Forum on Science Learning and Teaching*, 11(1), 1–19.
- Goldin, G. A. (2014). Mathematical representations. In S. Lerman (Ed.), *Encyclopedia of Mathematics Education* (pp. 409–413). Springer Netherlands. <https://doi.org/10.1007/978-94-007-4978-8>
- Goldin, G. A., & Kaput, J. J. (1996). A joint perspective on the idea of representation in learning and doing mathematics. In L. Steffe, P. Nesher, P. Cobb, G. Goldin, & B. Greer (Eds.), *Theories of mathematical learning* (pp. 397–430). Lawrence Erlbaum.
- Hahn, L., & Klein, P. (2022). Eye tracking in physics education research: A systematic literature review. *Physical Review Physics Education Research*, 18(1), 013102. <https://doi.org/10.1103/PhysRevPhysEducRes.18.013102>
- Hegarty, M., Mayer, R. E., & Green, C. E. (1992). Comprehension of arithmetic word problems: Evidence from students' eye fixations. *Journal of Educational Psychology*, 84(1), 76–84. <https://doi.org/10.1037/0022-0663.84.1.76>
- Hegarty, M., Mayer, R. E., & Monk, C. A. (1995). Comprehension of arithmetic word problems: A comparison of successful and unsuccessful problem solvers. *Journal of Educational Psychology*, 87(1), 18–32. <https://doi.org/10.1037/0022-0663.87.1.18>
- Höhne, J. K., & Lenzner, T. (2015). Investigating response order effects in web surveys using eye tracking. *Psihologija*, 48(4), 361–377. <https://doi.org/10.2298/PSI1504361H>
- 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 University Press.
- Huh, D., Kim, J., & Jo, I. (2019). A novel method to monitoring changes in cognitive load in video-based learning. *Journal of Computer Assisted Learning*, 35(6), 721–730. <https://doi.org/10.1111/jcal.12378>
- Joseph, A. W., & Muruges, R. (2020). Potential Eye Tracking Metrics and indicators to measure cognitive load in human-computer Interaction Research. *Journal of Scientific Research*, 64(01), 168–175. <https://doi.org/10.37398/JSR.2020.640137>

- 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](https://doi.org/10.1016/0010-0285(76)90015-3)
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329–354. <https://doi.org/10.1037/0033-295X.87.4.329>
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Koç-Şanlı, K., & Işık, C. (2020). Tam sayıların öğretim sürecinin öğretmenlerin model kullanımları üzerinden analizi [Analyzing the teaching process of integers through teachers' use of models]. *Bayburt Eğitim Fakültesi Dergisi*, 15(29), 81–108. <https://doi.org/10.35675/befdergi.519552>
- Kohout, S., Kruikeimer, S., & Bakker, B. N. (2023). May I have your attention, please? An eye tracking study on emotional social media comments. *Computers in Human Behavior*, 139, 107495. <https://doi.org/10.1016/j.chb.2022.107495>
- 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 technology in exploring learning from 2000 to 2012. *Educational Research Review*, 10(88), 90–115. <https://doi.org/10.1016/j.edurev.2013.10.001>
- Lee, W. K., & Wu, C. J. (2018). Eye movements in integrating geometric text and figure: Scanpaths and given-new effects. *International Journal of Science and Mathematics Education*, 16(4), 699–714. <https://doi.org/10.1007/s10763-016-9790-2>
- Lesh, R., Post, T., & Behr, M. (1987). Representations and translations among representations in mathematics learning and problem solving. In C. Janvier (Ed.), *Problems of representations in the teaching and learning of Mathematics* (pp. 33–40). Lawrence Erlbaum.
- Liang, C. P., & She, H. C. (2021). Investigate the effectiveness of single and multiple representational scaffolds on mathematics problem solving: Evidence from eye movements. *Interactive Learning Environments*, 31(6), 3882–3897. <https://doi.org/10.1080/10494820.2021.1943692>
- 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.
- Lin, J. J. H. (2024). AI-assisted evaluation of problem-solving performance using eye movement and handwriting. *Journal of Research on Technology in Education*, 1–25. <https://doi.org/10.1080/15391523.2024.2339474>
- Lin, J. J. H., & Lin, S. S. J. (2014). Tracking eye movements when solving geometry problems with handwriting devices. *Journal of Eye Movement Research*, 7(1). <https://doi.org/10.16910/jemr.7.1.2>
- Livingston, L. A., Sandals, L. H., & Vickers, J. N. (1992). Monitoring the effect of color on performance in an instructional gaming environment through an analysis of eye movement behaviors. *Journal of Research on Computing in Education*, 25(2), 233–242. <https://doi.org/10.1080/08886504.1992.10782046>
- Malone, S., Altmeyer, K., Vogel, M., & Brünken, R. (2020). Homogeneous and heterogeneous multiple representations in equation-solving problems: An eye-tracking study. *Journal of Computer Assisted Learning*, 36(6), 781–798. <https://doi.org/10.1111/jcal.12426>
- Mayer, R. E. (2009). *Multimedia learning*. Cambridge University Press.
- Miller-Singley, A. T., & Bunge, S. A. (2018). Eye gaze patterns reveal how we reason about fractions. *Thinking & Reasoning*, 24(4), 445–468. <https://doi.org/10.1080/13546783.2017.1417909>
- NCTM. (2000). *Principles and standards for school mathematics*. National Council of Teachers of Mathematics.
- Norqvist, M., Jonsson, B., Lithner, J., Qwillbaird, T., & Holm, L. (2019). Investigating algorithmic and creative reasoning strategies by eye tracking. *The Journal of Mathematical Behavior*, 55. <https://doi.org/10.1016/j.jmathb.2019.03.008>
- Obersteiner, A., & Tumpek, C. (2016). Measuring fraction comparison strategies with eye-tracking. *Zdm*, 48, 255–266. <https://doi.org/10.1007/S11858-015-0742-Z>
- Ott, N., Brünken, R., Vogel, M., & Malone, S. (2018). Multiple symbolic representations: The combination of formula and text supports problem solving in the mathematical field of propositional logic. *Learning and Instruction*, 58, 88–105. <https://doi.org/10.1016/j.learninstruc.2018.04.010>
- Paivio, A. (1990). Dual coding theory. In A. Paivio (Ed.), *Mental Representations* (pp. 53–83). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195066661.003.0004>
- Pallant, J. (2020). *SPSS Survival Manual: A step by step guide to data analysis using IBM SPSS* (7th ed). Routledge. <https://doi.org/10.4324/9781003117452>
- Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology*, 9(4), 441–474. [https://doi.org/10.1016/0010-0285\(77\)90016-0](https://doi.org/10.1016/0010-0285(77)90016-0)

- Plass, J. L., Moreno, R., & Brünken, R. (2010). *Cognitive load theory*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511844744>
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422.
- Rayner, K. (2009). Eye movements in reading: Models and data. *Journal of Eye Movement Research*, 2(5), 1–10.
- Schindler, M., & Lilienthal, A. J. (2019). Domain-specific interpretation of eye tracking data: Towards a refined use of the eye-mind hypothesis for the field of geometry. *Educational Studies in Mathematics*, 101(1), 123–139. <https://doi.org/10.1007/s10649-019-9878-z>
- Steffkovics, Á., & Kmetty, Z. (2022). A comparison of question order effects on item-by-item and grid formats: Visual layout matters. *Measurement Instruments for the Social Sciences*, 4(1), 8. <https://doi.org/10.1186/s42409-022-00036-z>
- 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(2), 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 and Review*, 18(3), 557–563. <https://doi.org/10.3758/S13423-011-0081-1>
- Susac, A., Bubic, A., Kaponja, J., Planinic, M., & Palmovic, M. (2014). Eye movements reveal students' strategies in simple equation solving. *International Journal of Science and Mathematics Education*, 12(3), 555–577. <https://doi.org/10.1007/S10763-014-9514-4>
- Susac, A., Bubic, A., Martinjak, P., Planinic, M., & Palmovic, M. (2017). Graphical representations of data improve student understanding of measurement and uncertainty: An eye-tracking study. *Physical Review Physics Education Research*, 13(2), 020125. <https://doi.org/10.1103/PhysRevPhysEducRes.13.020125>
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
- TMNE (2018). *Matematik dersi öğretim programı (İlkokul ve ortaokul 1, 2, 3, 4, 5, 6, 7 ve 8. sınıflar)*. MEB.
- von Glasersfeld, E. (1987). Learning as a constructive activity. In C. Janvier (Ed.), *Problems of representation in the teaching and learning of Mathematics* (pp. 3–17). Lawrence Erlbaum.
- Wickens, C. (2021). Attention: Theory, principles, models and applications. *International Journal of Human–Computer Interaction*, 37(5), 403–417. <https://doi.org/10.1080/10447318.2021.1874741>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Hilmi Karaca¹  · Erhan Ertekin² · Kursat Cagiltay³

✉ Hilmi Karaca
hilmikaraca@aksaray.edu.tr

Erhan Ertekin
eertekin@erbakan.edu.tr

Kursat Cagiltay
cagiltay@gmail.com

¹ Aksaray University, Aksaray, Türkiye

² Necmettin Erbakan University, Konya, Türkiye

³ Faculty of Education, Faculty of Engineering and Natural Sciences, TED University, Ankara Türkiye & Sabanci University, Istanbul, Türkiye