READING MATHEMATICS REPRESENTATIONS: AN EYE-TRACKING STUDY

Published version:

DOI:10.1007/s10763-013-9484-y

(Article begins on next page)

2015. READING MATHEMATICS REPRESENTATIONS: AN EYE-TRACKING STUDY. DOI:10.1007/s10763-013-9484-y. pp.1-23. In INTERNATIONAL JOURNAL OF SCIENCE AND MATHEMATICAL EDUCATION - ISSN:1571-0068 vol. 13 (S2)

Andrá, Chiara; Lindström, Paulina; Arzarello, Ferdinando; Holmqvist, Kenneth; Robutti, Ornella; Sabena, Cristina

Reading mathematics representations: An eye tracking study

Chiara Andrá[1], Paulina Lindström[2], Ferdinando Arzarello[1], Kenneth Holmqvist[2], Ornella Robutti[1], Cristina Sabena[1]

[1]University of Torino, [2]Lund University

e-mail: chiara.andra@gmail.com
phone: +39 0116702804
fax: +39 0116702878

Abstract. *We use eye-tracking as a method to examine how different mathematical representations of the same mathematical object are attended to by students. The results show that there is a meaningful difference in the eye movements between formulas and graphs. This difference can be understood in terms of the cultural and social shaping of human perception, as well as in terms of differences between the symbolic and the graphical registers, as they have been examined in literature. The results are also discussed in terms of didactic implications to support teachers in helping students to both deal with and to integrate multiple mathematical representations as well as acknowledge their own specificity.*
Keywords: visual perception, eye-tracking, mathematical representations, semiotics.

It is well acknowledged that reading a mathematical text requires giving sense to a set of multiple representations of mathematical objects: formulas, graphs, drawings and words. However, remedial literacy education on *text reading* has shown that the processes of "decoding" (namely, to recognize printed words) and "comprehension" (to understand what words means) are not highly correlated (see e.g. Lindamood, Bell & Lindamood, 1997). This urges us to briefly sketch theories that can help us defining what do we mean for "reading" and what do we mean for "understanding", and the relationships between them, on the one side, but on the other side it is necessary also to explicit the philosophical perspective from which we interpret such theories. Only after this necessary background, we will be able to position our research and point out the motivation and the relevance of our specific study. The structure of the first section, which serves as introduction, reflects this aim. The paper will continue with a literature review on researches related specifically to reading mathematical texts, and thus our research will be presented and discussed.

Reading and comprehension: a general 'ouverture'

Cognitive psychologists have been interested for decades about text processing in general: Kintsch (1998) argues that text comprehension involves processing at different levels, from the linguistic one, which entails decoding symbols, to the semantic analysis, which determines the meaning of what is read, forming propositional units according to the syntactic relationships and the coherence of the text, and thus inserting the text in the (proper) situation model, integrating information from the text with relevant prior knowledge. From the (basic) process of decoding signs to high-level modeling, Johnson-Laird (1983) maintains that humans rely on mental models of the world, and even perception yields a mental model. A mental model has the same structure of the situation it represents, but it is also partial, since it represents only certain aspects of it. Imagery is a fundamental aspect of the theory of mental models. Paivio (1971) has elaborated the Dual Coding Theory, which asserts that two distinct subsystems (a verbal one, specialized in dealing with language, and imagery, specialized for nonlinguistic objects and events) are activated when one recognizes, manipulate or thinks about words or things. In reading, when the subject combines pictures, mental imagery, and verbal elaboration, he understands and learns more effectively than just looking at plain text. These theories from cognitive psychology point to the complex relationship between reading a text and comprehending it. In our paper we investigate in particular how students *read mathematical texts*—meant as part of the cognitive process of understanding a mathematical text through its representations and thus having access to meaning.

With this respect, the eye-tracker as a method of investigation may provide useful insights. Radford (2010), in facts, examining how the eye of the student is d*omesticated*, namely how it is culturally educated to frame the perception of mathematical objects in particular, cultural ways, arrives to the conclusion that the relationship between our senses and our understandings is historically shaped in certain ways as people engage in social and cultural practices. Moreover, it is the same sensorial perception that is not a purely physiological substratum, but it is understood as a social process. Thus, also the visual perception is framed in a way such that “what we see is not the result of direct inputs but of stimuli already filtered by meanings and information about objects and events in the world – meanings conveyed by language and other semiotic systems” (Radford, 2010, p. 2). Following an empiricist tradition (see de Freitas & Sinclair, 2012), we see that perceptual routine habits and material interactions constitute conceptual categories. Thus, focusing on the movements of the eye when a student is reading a mathematical text can provide us with insights on his learning process: the material, imagery or verbal, stimulus is taken by the visual perception, which entails a situational model that allows the reader to read what is in front of his eyes. Visual perception is thus educated, and the product of such education is the mental models that allow the individual to see the world. Our focus is on different ways of reading different kinds of mathematical representations. Within this very general claim, we address the issue by means of an explorative study aimed at investigating what can (and what cannot) be “seen” by using an eye-tracker, and thus what we, as mathematics education researchers, can (and what we cannot) learn from it, or at least what other fields of knowledge as well as results within the mathematics education scientific community can be confirmed, and to what extent.

We focus our study on the reading of formulas, graphs, and plain texts in mathematics. The relevance of this choice is discussed in the following literature review. We use eye tracking methodology to analyze the reading of the aforementioned representations: the variables (fixation duration, number of fixations, dwell time, pair-wise comparisons vs. overview looking –see Holmqvist *et al.*, 2011) are presented in the methodology section. Are there differences between formulas and graphs, in terms of the movements of the eye? What do such differences, if any, entail in terms of reading and understanding a mathematical text? After having shown the results from our experimental study, in the discussion we attempt an answer to our research question, and we sketch possible future directions.

Reading mathematics

Vision has its roots in both biological and socio-cultural history of mankind: thus, we are involved in seeing both what comes to be seen and what we are unable to see, but we maybe aspire to see (McCormick *et al.,* 1987). This duality is especially present in mathematics, and consists both in the objects one looks at, and in the meanings they have as abstract objects in mathematics.

Moreover, mathematical objects are cultural and sophisticated products, and they are accessible not in a direct way, but only by means of *representations* (e.g. Duval, 2006): this feature of mathematics has yielded to state that there is no *noesis* without *semiosis* (Duval, 2006), namely that cognitive acts such as understanding a concept are not possible without the coming into being of a relation between the signifier (a sign, a representation) and the signified (the mathematical object), as it has also put forward by Johnson-Laird (1983) referring to Peirce. The mutually constitutive nature of sign and object in mathematics is also central in Sfard's (2000) work about the discursive element of learning mathematics. Moreover, the sign system(s) of mathematics constitute a referentially closed world: in this sense, taking a semiotic lens for examining mathematics learning is a privileged point of view, given the highly structured semiotic systems that are central to the mathematical activity (Leung, Graf, & Lopez-Real, 2006). On the counterpart, it is not enough to stay in front of a visual stimulus in order to understand its meaning: perception is *inadequate* (Levinas, 1989), since it is selective or *intentional* (Husserl, 1931). In the learning processes, the main issue for the students becomes looking at representations in a certain, intentional and cultural way (Radford, 2010). Mathematical texts are complex, in that they include formulas, graphs, numerical tables, natural language and so on, and each representation must be properly interpreted, and coordinated with the others, thus complementing, constraining, and supporting the construction of deeper knowledge (Ainsworth, 2006). Moreover, the reading may be influenced by a variety of factors, such as one's knowledge, the context of the task, the structure of the representation itself, the mental models activated by students (Johnson-Laird, 1983), how representation is encoded and the role of verbal-imagery components (Paivio, 1971), how text is decoded and comprehended at various levels (Kintsch, 1998), and so on.

Kellman, Massey and Son (2010) have shown that *perceiving what is important to attend to* enhances students' ability to correctly and faster determine for example which variable to look for in an equation when determining its slope.

Andrà *et al.* (2009) considered two groups of students: one corresponds to mathematically educated students, which are referred to as experts, the other one corresponds to a low level in mathematics. Novices' versus experts' ways of reading different mathematical stimuli have been compared, and it has been found that experts' eye movements have a coherent and synchronic mode of attending to mathematical stimuli, while novices' eye movements follow diversified paths. This result can be interpreted as a common, consistent and cultural way to attend to mathematical stimuli by mathematically educated students, whilst for novices the mathematical representation may be navigated in various, non-methodical manners. Similar differences between experts and novices have also been observed by Inglis & Alcock (2012), and Eppelboim & Suppes (2001).

Cognitive load theory is concerned with the ways instructional materials can enhance learning, or on the opposite impede it. Tarmizi and Sweller (1988), in the learning of geometry, found a split attention effect showing that an integrated representation enhanced learning compared to a split representation where students instead had to map formula information on to the geometric figure. The integrated format of the task reduced students´ cognitive load while learning. Chandler and Sweller (1991) conclude that the integration of multiple source of information is beneficial when they are unintelligible unless integrated. This is in accordance with the aforementioned semiotic approach, in that the coordination of multiple representations in more than one semiotic register is fundamental in mathematics learning (Duval, 2006), and in text comprehension in general (Paivio, 1971).

The aim of our paper is investigating *how the mathematical representation seen as a cultural product reveals itself to the learner*, and two cognitive processes turn out to be important when focusing on the visual experience of the learning subject: the first one is concerned with the process of attending the representation in order to access its meaning, and the second one regards the shift from one representation, to another one, of the same mathematical object (Duval, 2006). Namely, how a mathematical representation let itself to be attended to by the learner, and how it reveals the same mathematical object represented in another semiotic register. For example, how a parabola reveals itself by means of a formula and how this is attended to by the learner's eye when it is represented in a graphical register. We would like to stress, at this point, that we are situating the study within a cultural perspective, since mathematical objects are a product of our culture, but we are not going to make a cultural comparison between for example how students in different countries read mathematics or how instructional practices influence how students attend to math representations.

In order to address the issue regarding the way(s) the mathematical representation as a cultural product reveals itself to the learner, we consider a methodology of research that makes use of the eye-tracker. The eye-tracker allows us to track the movements of the eye while a participant is performing a task: it consists of a camera, which is filming the movements of the eye, which can be visualized in different forms. Results in Neurosciences have shown that human cognition is embodied: not only the areas in the brain that are delegated to high mental processes are proximal to the ones related to sensorial experiences, body movements, and actions, but mostly cognition is strictly grounded in the human body, and in its location in space and time (Lakoff & Johnson, 1980; Lakoff & Núñez, 2000; Seitz, 2000; Gibson, 2002). In that, eye-movement is part of sensorial experience, and Radford's understanding of human senses brings us to claim that *investigating the relationship between the mathematical representation and the movements of the eye may shed light on the way human beings access the mathematical knowledge*. Researches in eye-tracking have shown that there is a correlation between what is looked at and what is being attended to (Rayner, 1998; Just & Carpenter, 1980; Yarbus, 1967; Bushwell, 1935). These results agree with other researches supporting a stronger correlation between fixations and cognitive processing of the information (Latour, 1962; Volkman, 1976). Eye tracking is being used more and more within educational research (Scheiter & van Gog, 2009; van Gog & Scheiter, 2010). The merit from a didactic perspective is that we can examine how and which information students are attending to.

Formulas, graphs and text: how meaning is conveyed

Plain text is always present in mathematical lessons, books, and journals. It is used both as a fundamental means to communicate, and as a means to express mathematical problems (Duval, 2006). But plain text is also the mode of expression of other activities that lie at the core of mathematics, such as exploring, conjecturing, problem solving, arguing, and proving. Kintsch (1998) argues that (text) comprehension complements but by no means supplants problem solving. In agreement with this study, we focus our attention on text *reading* in terms of the movement of the eye. But plain text is only part of mathematical realm: formulas and graphs constitute other important representations. What does it mean *to read a formula*? The use of ***formulas*** and specific symbols, together with their meanings, can be framed within Arcavi's (1994) expression *symbol sense*, which indicates "a complex and multifaceted 'feel' for symbols […] a quick or accurate appreciation, understanding, or instinct regarding symbols" (Arcavi, 1994, p. 31). Other authors underlined that mathematical symbols are at the same time "processes" and "objects", and that such dual nature is difficult to manage by students (Sfard, 1991). To highlight the amalgam of these two natures (process and object) within symbols, Gray and Tall (1994) spoke of "procepts". They define an elementary procept as the amalgam of three components: a process, a mathematical object produced by the process, and a symbol which is used to represent either process or object.

In the passage from text to formula, a *conversion* from one semiotic register to another one in Duval's (2006) words, there is a *semiotic contraction*: the learner comes to recognize and attend to the essential elements within an evolving mathematical experience (Radford, 2002). A formula constitutes both a procedure and a *symbolic narrative*, "recounting, with symbols, a previously linguistically objectified schema" (Radford, 2002, p. 18). We can say that formulas require to pack the information contained in the plain text, and to do it symbolically and according to formal rules. In fact, each symbol derives its function in the formula according to its position (e.g. in $2x+x^2$ the symbol '2' is the coefficient of x in '2x' and it is the exponent in 'x^2'), and formulas often have a non-linear structure (e.g. the structure of the formula 3x + 2(x + 5y) does not follow the sequential order commonly used when reading a text: the straightforward reading "three times x plus two times x plus five times y" could be misleading, since it could refer to another structure, that is 3x+2x+5y). A formula is supposed to be read according to the rules given by its semiotic register (Duval, 2006). Thence, in the opposite direction, the conversion from formula to plain text requires an un-packing of such symbolic information, in order to give a narrative and discursive understanding of it. From this exposition emerges that the theoretical framework provided by Kintsch (1998) on textual comprehension at different processing levels applies to comprehension of mathematical formulas as well: decoding symbols, recognizing their role within the syntax of the formula, perceiving the global sense of the formula out of any single symbol, and situating it with a proper model. Specifically, in our study we look at formula-plain text conversions in both directions. As Radford stressed in several works (see e.g. Radford, 2002; 2010), a crucial problem for learners is to designate in different ways the objects of discourse, in the two semiotic registers (this has also been pointed out by Duval, 2006). On the one hand, the natural language plays a prominent role in the signs-making-sense process when dealing with (new) symbolic systems (Radford, 2002). On the other hand, the letters in formulas are not mere substitutes of nouns, but the mode of designation of the objects of discourse is functional (Radford, 2002).

These theoretical premises not only lead us to claim that the conversion from text to formula, and from formula to text, is a central cognitive act in the learning process of those areas of mathematics where symbols play a significant role, such as algebra, but also urge us on investigating which cognitive demand is beyond formulas. Following also the aforementioned works by Sfard (1991), and Gray & Tall (1994), on the procedural nature of formulas, we highlight that, when dealing with a formula, students are often expected to manipulate it: as an example, we just quote the work by Kieran (1988), which considers the procedural meaning associated to the equal sign by the students, as if formulas always need to be transformed *in something else* within the same symbolic register. This "manipulative"-demanding feature of formulas does not seem to belong to graphs in mathematics, where the students are often expected to rather interpret them. Graphs can be meant as other kinds of modeling the world (Johnson-Laird, 1983) in mathematical terms. Whilst in formulas the verbal component is dominant, in graphs it is the imagery one (Paivio, 1971).

A ***graph*** usually contains information that is spatially spread out and has many iconic aspects (Tall, 2002). Graphs are structured in holistic blocks that require a global reading (and not subsequent un-packaging as per formulas). In addition, graphs have an iconic component that is absent in formulas (like the increasing of a graph, its being bigger/smaller than, etc.). Learning how to interpret a graph, a student has to integrate all the elements of the graph into a comprehensible whole (Robutti, 2006). If this is not taught, students risk to focus only on certain parts of the graph, without acknowledging that there may be useful and necessary information elsewhere in the representation. In order to make sense of a graph, a suitably strategy can be to make a narrative by linking the shape of the graph to the phenomenon it is modelling. An interesting example about modelling motion can be found in Nemirovsky & Monk (2000). By means of the narrative, different elements of the graph are integrated into a coherent whole. In our study we will see how different graphs are converted into proper plain texts. Similarly to formulas, in fact, reading a graph requires a subject to locate and decode where the relevant information is (e.g. Carpenter & Shah, 1998). A graph usually condenses a lot of information, which is not always explicit. Again, Kintsch's (1998) distinction between decoding and comprehending a text/graph supports our claim that a graph in mathematics should be read in specific ways, decoding relevant information, finding relationships at the level of syntax, discerning the overall meaning of the graph, and positioning it with respect to the situation it models.

The experimental design and methodology

According to the literature review already presented, our research problem addresses the issue of giving sense to representations of mathematical objects in the registers of natural language, symbols and graphs. The same mathematical object can be represented in one of the three registers, and among all the possible conversions from one semiotic register to another one (Duval, 2006), we select three that seem to be crucial to the sense-making process in mathematics: from formulas to plain texts, from plain texts to formulas, and from graphs to plain texts. Also other conversions are crucial in mathematics: from text to graph, and from formula to graph and viceversa. In this exploratory study, however, without discarding the importance of the latter group of

conversions, we focused on the former, where plain text is understood as the verbal component which carries the meaning of the mathematical object, as it is 'seen' by the students. We designed 43 stimuli, given to the students in random order for avoiding fatigue effect, of three types: (1) Formula-text (ft type, N=15 items), the input is a formula with four text alternatives; (2) Graph-text (gt type, N=12 items), the input is a graph with four text alternatives; (3) Text-formula (tf type, N=16 items): the input is a text with four formula alternatives. The task was to determine which of the four alternatives is an accurate representation of the input, or corresponds to the input according to some rule or feature. The correct answer was always present among the alternatives.
We decided to not let the participants use paper and pencil, though we are aware that they are helpful tools when solving mathematical problems. In fact, we focused this study on the eye movements rather than on the construction of the answer. In a future study we plan to look at both the reading of the representation as well as the writing process when students are constructing an answer.
We outline that our focus is not on the *difficulty* the students have in switching from formulas or graphs to texts in mathematics, but on the way the stimulus is navigated by the eye of the students. For this purpose, we make use of a common concept used in eye tracking methodology that allows the identification of those parts of the stimulus that contain relevant information: the *areas of interest* (Knoblich *et al.*, 2001; Verschaffel *et al.*, 1992; Hegarty *et al.*, 1992). By using an area of interest with a semantic mathematical content we can say something specific about how information in that area is processed. For each stimulus we consider five areas of interest: the input, and each of the 4 answers. The example in figure 1 shows a stimulus that requires a conversion from a formula to plain text: in fact, the input contains a formula and the four alternatives contain plain text. One can notice that plain text may contain some symbolic expression, but the overarching cognitive process is to pass from the formula to plain text. Since our interest is on the conversion from one semiotic register to another one, we try to avoid the presence of more than one register in the same area (both the input and each alternative), as much as possible.
In the eye-tracking methodology, it is possible to use the areas of interest in different ways depending on how fine-grained the analysis is. In fact, there are multiple levels at which areas of interest can be implemented: *macro*level (e.g. the 5 big areas: input and 4 alternatives), *meso*level (e.g. within each of the 5 areas), or *micro*level (e.g. examining a specific part of the graph, or formula, or text such as a slope, or 2x, or a specific word). For more details we refer the reader to Holmqvist *et al.* (2011). Since the purpose of our study is to understand how the representation *as a whole* reveals itself to the learning subject, and how the connection between two different representations of the same mathematical object is established, in this paper we operate at the *macro*level, considering the area of interest as the area of the whole input/alternative. We use areas of interest to examine "overall behaviour" since we are not looking at the detailed behaviour of every single transition in this study, but this does not mean that it is not possible to look at each transition in more detail. If we, instead, would focus our study on one specific stimulus type and/or a specific problem, such as time-motion relations or solving a specific equation, it would be interesting to do a more detailed analysis and for example compare it with different solution strategies. For example, this is what the match stick problems/insight studies (Knoblich, Ohlsson, & Raney, 2001) have done and also it is shown in word problems that certain areas are useful to attend to in order

to solve them (Hegarty, Mayer, & Green, 1992; Hegarty, Mayer, & Monk, 1995; Verschaffel, DeCorte, & Pauwels, 1992).

We further stress that our focus is on how the mathematical object reveals itself by means of different representations (formula and text in this example), and how the subject makes sense of the representation by selecting one out of four alternatives (in the text register in this example), rather than in reading or navigating with the eye a certain area of the stimulus.

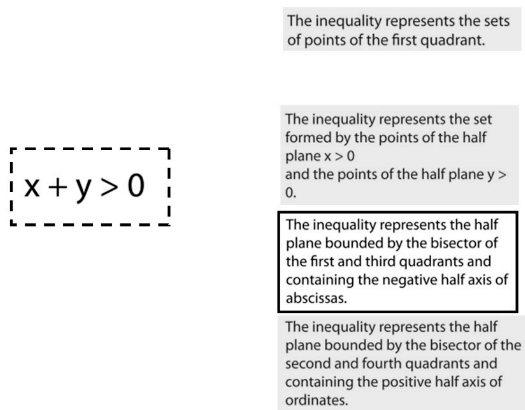

Figure 1. Five areas of interest. The dotted box corresponds to the input. The black box corresponds to the correct alternative and the grey boxes refer to the incorrect alternatives.

All the 43 stimuli can be found on the Electronic Supplementary Material. Here, we would like to draw the reader's attention to the fact that the stimuli were different not only in type (ft, gt or tf), but also in complexity and difficulty *within the same type*. In general, since we are interested in the sense-giving activity as far as sense is conveyed by natural language in plain text, we did not want too easy or too evident alternatives for the stimuli. Instead, we did want to have some reasoning behind the process that brings a subject to choose one alternative or another. Part of this process consists of navigating the area of interest with the eyes: according to the semiotic register, such navigation can differ. Since navigation is not mere sensorial perception, but is shaped by culture and social interaction, we claim that the differences in eye navigation between different semiotic registers can shed light on how such representations are dealt with by the subject. The question is: given the diversity of the stimuli in both complexity and difficulty, is it possible to detect such differences in attending different semiotic registers? We will use eye-tracking methodology, which allows us to consider variables that may vary according to the different type of stimuli.

46 Swedish university students were recruited based on their different knowledge backgrounds in mathematics. 24 had no previous university studies in mathematics. 22 had one year of studies in a faculty of engineering. The differences in performance according to the mathematical background are shown in Andrà *et al.* (2009). As stated earlier, in this study we concentrate only on the differences between representation types of the stimuli. This is also in accordance with the question that ends the previous paragraph: Our aim is to detect (if any) differences in the type of stimuli and in the kind of representation, *despite* the variability in item difficulty, complexity, as well as in the participants' background. As a consequence, we are not classifying the stimuli in terms

of difficulty level and/or complexity; as well, a literature review on these issues is not in line with the interest of our research, and thus is not present. We, instead, would rely on statistical analysis to answer the following research question: is there a *commonality* that the stimuli of the same type share? A positive answer to this question, provided by an *a posteriori* statistical analysis, may contribute to the research in the field, for a better understanding of how the students deal with semiotic representations.

Now, we present and discuss the variables, from the eye-tracker methodology, taken into account in this study. The eye-tracking methodology provides us with variables that can be considered as indicators of reading behavior: average fixation duration on the input area of interest, number of fixations on the input area of interest, dwell time on the input area of interest, and types of transitions on the stimulus (navigation on the whole stimulus): pair-wise comparisons vs. overview looking (Holmqvist *et al.*, 2011). Fixation *per se* refers to the time when the eye is still on a certain area of the stimulus. Fixation duration is regarded as a measure of processing capacity. In our study fixation duration is the average of all fixations in a given area. If we have 2 fixations of, respectively, 200 and 300 ms, then the average fixation duration is 250 ms. A long fixation on a word in reading usually means that the word is difficult to interpret (Rayner, 1998). However, in scene perception studies, a long fixation often means that something interesting is fixated (Henderson, Weeks, & Hollingworth, 1999).

A difference in average fixation duration on the input area with respect to the kind of stimulus (ft, gt, and tf) may reveal how long the input area requires to be fixated in order to establish a connection with the alternative *on average*. Our research hypothesis in the current study is that fixation durations for formulas are the longest, followed by graphs, and the shortest for text. In fact, the imagery component in graphs, as well as the verbal one in plain text, may foster the meaning-making process in a faster way with respect to formulas, which entail both a figurative/symbolic and a textual components that both need to be unpacked in a cultural, sophisticated way, thus requiring more time to be processed.

The number of fixations can indicate how certain content in an area of interest is processed. If an area receives a high number of fixations it can mean that the information is dense or complex and therefore needs to be re-examined multiple times. For example, a word that has many fixations is usually a difficult word (Rayner, 1998). It has also been shown that semantically important information increases the number of fixations (Henderson, Weeks, & Hollingworth, 1999; Loftus & Macworth, 1978). A difference in the number of fixations on the input area with respect to the type of stimulus may reveal whether one type of stimulus is looked at many or few times. Combining this information with the fixation duration, we can get an idea of whether a kind of stimulus is fixed few times but longer, or many times but shorter, than another one. Following this, we predict the most number of fixations for formulas, followed by graphs and the shortest for text.

Dwell time is the sum of all fixations on a given area. For example if we have 2 fixations with durations of, respectively, 200 and 300 ms in an area, then the dwell time is 500 ms. A longer dwell time is expected when an area contains complex information because more cognitive processing is required to interpret it. For example, the time it takes to interpret a graph is highly related to the number of unique qualitative relations in the graphs as shown by Carpenter and Shah (1998), and as such a longer dwell time indicates that there is more information to process. A shorter dwell time on the formula input area with respect to the graph may reveal that the subjects identify the relevant information in the formula more quickly than in the graph, while a long dwell time on the text input area may just confirm that text requires some time to be read, being made of several words. A comparison among the three kinds of stimuli may reveal how long each one is attended, beyond its structure and its features.

The three measures described take into account the area of the input. This area, in fact, requires particular attention since it contains, in two out of three types of stimuli, the semiotic representation of the mathematical object that needs to be matched with its textual description. In order to investigate the sense-giving process understood as establishing a connection between text (the realm of natural language) and another representation, we consider the number of transitions between different areas of interest. The transitions are considered to correspond to conversions between different registers. Since this is an exploratory study on a macro level with respect to the area of interest, we further divide and examine in more detail two types of transitions. The first is pair-wise comparison, which refers to transitions between two specific areas of interest: the participant in this case goes back and forth between the same two areas of the stimulus for a number of times. The second is overview looking and refers to scanning of all areas: the participant navigates the whole stimulus, focusing on almost all the areas in it. It has to be said that pair-wise comparison and overview looking are only two cases among all the possible transitions in an eye tracking study. In our research, we take into account only these two behaviours since we assume that they are the most interesting from a didactical and cognitive point of view: pair-wise comparison refers to comparing two representations of mathematical objects, and overview looking refers to grasping the overarching meaning. In a sense, pair-wise comparison can be traced back to Kintsch's (1998) text processing at microlevel, the search for and the analysis of coherence relations among different areas of the stimulus; overview looking can be meant as the macrostructure, which entails also the recognition of global topics and their interrelationships (Kintsch, 1998). According to Kintsch, micro- and macro-structure together form the textbase, which represents the meaning of the text. Hence, looking at pair-wise comparisons and overview looking is important in a study using the method of the eye-tracker, since these two kinds of eye-movement parallel two relevant cognitive processes. In order to quantify them, it is necessary to take into account sequences such as I-A-I-A-I-A-B-C-D-I-A-B-C-D. In the first part of the sequence, the student is performing a pair-wise comparison between the input (I) and the first alternative (A), while in the second part of the sequence she is navigating all the five areas (input and A, B, C, D alternatives).

Since our stimuli contain a correct alternative to be detected, we also consider the percentage of correct answers given by participants with respect to the kind of stimuli. We understand this variable as an indication of how the mathematical object reveals itself to the learning subject, in that a difference (if any) in the percentage of correct answers with respect to the kind of stimuli may contribute to examine the process of establishing a (correct) relationship between two semiotic representations of the same mathematical object. The presentation of the results begins with these percentages.

Results and analysis

Our data show that the highest percentage of correct answers are in the group of graph-text stimuli (gt-type): 0.52, with 95%-confidence interval between 0.48 and 0.56;

followed by the group of text-formula ones (tf-type): 0.46, with confidence interval 0.43-0.49; the lowest one regards the group of formula-text stimuli (ft-type): 0.33, with confidence interval 0.36-0.30. If we consider the percentage of correct answers as an indicator of how easily the mathematical object reveals itself to the learning subject, meant as the latter's ability to establish a correct relationship between two semiotic representations of the former, then we can say that when the input is a graph, it is more easy to find the textual correspondent for students. The most difficult gt-type stimuli turned out to be GT09, GT11, GT12, the easiest GT02, GT16, and GT18. If one looks closely at them (on the Electronic Supplementary Material), one concludes that there is not a strikingly common feature, in terms of eye movements: on the contrary, we can link the difference in correct answers percentage to the difference in the cognitive functioning. The items that *a posteriori* turned out to be the easiest are not "easier to be read", but easier to be dealt with in mathematical and didactical terms: they require commenting on the trend of the graph. Conversely, difficult gt-type stimuli require to make some computation, or to have some knowledge of calculus. This may help confirming our starting hypothesis in the subsequent analysis: some variables assume values that significantly differ from one type of stimuli to another one, even if the *individual differences* between two stimuli of the same type are almost significative according to other variables. The same can be said for the other two types of stimuli: for example, the ft-type most difficult stimuli (FT03, FT07, FT14, FT16) require some reasoning, while the easiest ones (FT02, FT04) require commenting on some features of the formula given in the input.

Bar-charts in Figure 2 show the average and the 95% confidence intervals for dwell time, fixation duration, and number of fixations. Since in almost all the cases they do not overlap, we infer that differences between types are almost always significant.

Observing Figure 2-left, the longest total dwell time is for the graph-input (11.2 seconds on average), followed by the text (7.9 s), then the formula (7 s). Since participants were given 40 s maximum for attending the whole stimulus, we can notice that the total dwell time on the input area is not less than 6 and not more than 12 seconds, which correspond to 15%, and 30% respectively, of the whole time per stimulus.

From figure 2-middle we observe that the longest fixation durations are found for the formula input and the shortest for both the graph and the text. We can notice that fixation duration *on average* is not shorter than 190 ms and not longer than 250 ms: the range is quite narrow (it is expressed in milliseconds!). Looking at the width of confidence interval, we can notice that the variability with respect to this variable is few (the width of the confidence interval is 2 to 4 milliseconds), namely we can suppose that all the stimuli of the same type have very similar fixation duration.

Looking at Figure 2-right, we notice that there is the greatest number of fixations when the input is a graph, followed by the text input, and by the formula input.

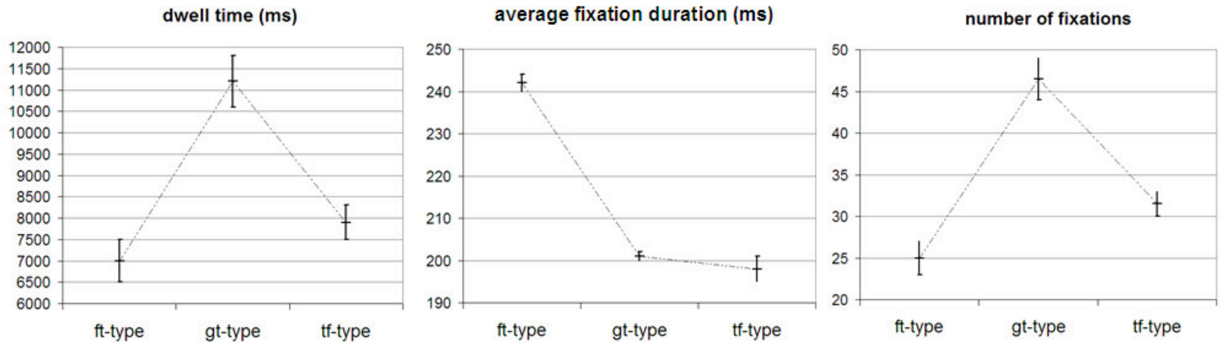

Figure 2. On different representation types: left) dwell time; middle) number of fixations; and right) average fixation duration. 95% confidence intervals are reported, and show significant differences between the estimated values.

We now examine the transitions in terms of pair-wise comparisons and overview looking. We recall that 'transitions' here are meant as the eye-movements that parallel conversions (Duval, 2006) at the cognitive level. The pair-wise comparisons can be within the same register (if they are both alternatives) or conversions between two different registers (if they are input-alternative). Figure 3-left shows that there are some differences between the case in which a graph is present with respect to formulas regarding the pair-wise comparisons: gt-type stimuli have lower pair-wise comparisons in percentage. In Figure 3-right, we see that there is more overview looking for the ft-type compared to the gt- and the tf-type. Looking at the percentages range, one notices that pair-wise comparisons and overview looking correspond to around 10% of the total transitions each. What about the remaining 80%? We have observed *a posteriori* that 70-75% of the total number of transitions regards eye-movement *within* the same area of interest at the *macro*level: the majority of transitions, in a sense, regards navigating the same area of interest (it is possible to "see" transitions within the same area of interest in the eye tracker method, if one further divides the area of interest in sub-parts, and considers transitions between such sub-parts, which can be as fine as the pixel definition of the screen). This is not surprising, given the kind of areas the students have to deal with. Moreover, considering data from dwell time above, we know that 15% to 30% is the proportion of time the students look at the input. Hence, the remaining 45%-60% is the proportion of time in which participants look at the four alternatives.

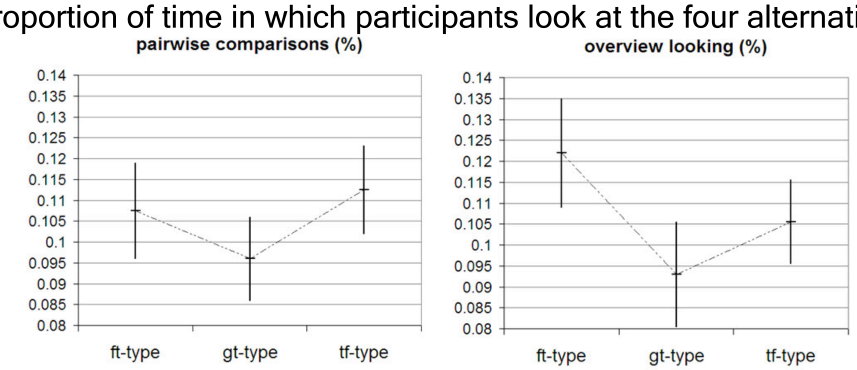

Figure 3. Percentages and confidence intervals of: left) pair-wise comparisons, and right) overview looking for the three kinds of stimuli.

To sum up, we have observed that, for gt-type stimuli, the area of the input is dwelled for the longest time and for the highest number of fixations, but the average fixation duration is the shortest. As a consequence, graphs in our study have been fixed many times, but for a short period of time. Both pair-wise comparisons and overview looking are fewer in percentages, hence we can draw the conclusion that the students were engaged in attending the graphical input area rather than comparing it with textual representations.

For ft-type stimuli it is the other way around: the shortest dwell time, the lowest number of fixations, and the highest fixation duration: when the input is a formula, it is fixated for longer time, but for a smaller number of times. Overview looking is the highest in percentage (with respect to the other two types), hence we can conclude that ft-type stimuli lead the students to attend the formula (they look at it for a relatively long time, but for few times), and to navigate the whole stimulus in order to find the correct textual correspondent. This is in accordance with the analytical nature of formulas. In looking at a formula, a subject needs fewer but longer fixations, since she focuses on the specific relevant elements (e.g. an exponent, a sign, a coefficient). Instead, a graph is usually perceived in a more holistic and condensed (synthetic) manner, which requires more but shorter fixations by the subjects.

As regards tf-type, some values are close to ft-type, others are different: dwell time is short, but not as short as the ft-type case; fixation duration is very low, and the number of fixations is in between gt- and ft-type. Pair-wise comparisons are slightly higher in percentage, while overview looking is lower, with respect to ft-type. When the input is text, and the alternatives are formulas, the students look quickly at the text (low fixation duration, quite low dwell time), and they look more at the alternatives: we draw this conclusion from the fact that the natural language of the text is more immediate than the symbolic language of formulas. Hence, the students make more pair-wise comparisons and attend less the textual input, since they focus on the alternatives, which are given in a symbolic and sophisticated language.

Discussion

The main result of the study consists in the difference between two representations that are fundamental in mathematics: formulas and graphs. In a single sentence, one can say that on average in reading *formulas* subjects have a *smaller number of longer fixations* than *graphs*, which present a *bigger number of shorter fixations*. Looking also at dwell time, we can see that there is not a balance between number of fixations and average fixation duration in formulas and graphs: dwell time, in fact, is higher for graphs, telling us that the big number of short fixations on the graph results in an overall longer dwell time on graphs with respect to formulas.

Formulas turned out to be the most difficult type of stimuli, since the percentage of correct answers was the lowest one. When the input is a formula, there is higher overview looking, which we have related to Kintsch's (1998) macrostructure processing level: the students seem to spend more time in navigating the global structure of the stimulus. We can interpret these results as follows: a formula condenses its information in a shorter but more hidden inscription. Both verbal and imagery components (Piavio, 1971) are present, but in a highly cultural (and by no means immediate to the reader) fashion. After a first overall glance, in order to understand the formula it is necessary to find out its structure: the rules for "reading" a formula are provided by its semiotic register (Duval, 2006). If a person is not confident with such rules, s/he hardly deals with

formulas: this can be a reason why the percentage of correct answers is low for ft-type stimuli. Moreover, the rather procedural nature of formulas (Gray & Tall, 1994) may lead the students to use automatic strategies and speedy elaborations, instead of interpreting the meaning of the given formula.
On the opposite, graphs turned out to be the easiest, and pair-wise comparisons as well as overview looking were relatively few (given that the majority of time is spent on the graph area of interest). However, we conjecture that this is not only a matter of time: the graphical representation does not encourage the students to read the textual counterparts many times (as it happens for formulas, which force more overview looking). Unlike formulas, graphs do not invite to automaticity in terms of using a given strategy of how to read it. Even though there are rules governing how a graph ought to be interpreted, the perceptual features of a graph (namely, its imagery component) can be more intuitive to grasp with respect to other kinds of representations, as suggested also by Paivio (1971).
When the input is text, dwell time and number of fixations have values that make text close to formulas (namely, quite low), but fixation duration is similar to graph (low). We can infer that text in the input is attended for shorter time, both in terms of duration and in terms of the number of times the students come to see it. Moreover, the relatively high number of pair-wise comparisons leads us to say that, when the alternatives are given in symbolic language, the students are fostered to compare the representations in pairs. In a sense, the tf-type does not merely represent an intermediate situation between ft- and gt-type (as it could be seen from results on dwell time, number of fixations, and overview looking): as for the symbolic and the graphical inputs, also the textual one is attended by the students in a specific way.
The results already summarized allow us answering affirmatively to our research question: there exist some common features (in terms of eye movements) within stimuli of the same type, despite the differences in both the subjects involved (for example, in terms of the mathematical background), and the stimuli themselves (in terms of difficulty, complexity, or cognitive demand in general). And the eye-tracker methodology provides us suitable tools in order to characterize such differences in terms of different ways of navigating the stimuli. The visual perception is understood not only as a physiological substratum, but mostly as culturally and socially shaped sense, which allows us attending mathematical representations in certain, structured way.

References

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*, 183-198.

Andrà, C., Arzarello, F., Ferrara, F., Holmqvist, K., Lindström, P., Robutti, O., and Sabena, C. (2009). How students read mathematical representations: an eye tracking study. *Proceedings of the XXXIII Conference of the Psychology of Math Education*. Thessaloniki, GR.

Arcavi, A. (1994). Symbol sense: Informal sense-making in formal mathematics. *For the Learning of Mathematics, 14*(3), 24-35.

Bushwell, G. T. (1935). *How people look at pictures.* Chicago: University of Chicago Press.

Carpenter, P. A., and Shah, P. (1998). A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied, vol. 4*(2), 75-100.

Chandler, P., and Sweller, J. (1991). Cognitive Load Theory and the Format of Instruction, *Cognition and Instruction, 8*(4), 293-332.

De Freitas, E., and Sinclair, N. (2012). Diagram, gesture, agency: theorizing embodiment in the mathematics classroom. *Educational Studies in Mathematics, 80*

Duval, R. (2006). A cognitive analysis of problems of comprehension in a learning of mathematics. *Educational Studies in Mathematics, 61*, 103-131.

Eppelboim, J, & Suppes, P. (2001). A model of eye movements and visual working memory during problem solving in geometry. *Vision Research, 41,* 1561-1574.

Gibson, J. J. (2002). A theory of direct visual perception. In A. Noe and E. Thompson (Eds.), *Vision and Mind* (pp. 77-110). Cambridge, MA: the MIT Press.

Gray, E. and Tall, D. (1994) Duality, Ambiguity, and Flexibility: A "Proceptual" View of Simple Arithmetic. *Journal for Research in Mathematics Education, 25*(2), 116-40.

Hegarty, M., Mayer, R. E., and Green, C. E. (1992). Comprehension of arithmetic word problems: Evidence from students' eye fixations. *Journal of Educational Psychology, 84*(1), 76-84.

Henderson, J. M., Weeks, P. A. Jr., and Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scence viewing. *Journal of Experimental Psychology: Human Perception and Performance, 25*(1), 210-228.

Holmqvist, K , Andrá, C., Lindström, P., Arzarello, F., Ferrara, F., Robutti, O., and Sabena C. (2011). A method for quantifying focussed versus overview behaviour in AOI sequences. *Behavior Research Method*, *43*(4), 987-998.

Husserl, E. (1931). *Ideas. General introduction to pure phenomenology.* (W.R. Boyce Gibson, trans., 3rd [1958] edn.), London, UK: George Allen and Unwin.

Inglis, M., and Alcock, L. (2012). Expert and Novice Approaches to Reading Mathematical Proofs. *Journal for Research in Mathematics Education, 43*, 358-390.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, interference and consciousness.* Cambridge, UK: Cambridge University Press.

Just, M. A. and Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review 87*(4), 329-354.

Kellman, P., Massey, C., & Son, J. (2010). Perceptual learning modules in mathematics: Enchaning student´s pattern recognition, structure extraction, and fluency. *Topics in Cognitive Science, 2,* 285-305.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* New York: Cambridge University Press.

Kieran C. (1988). Two different approaches among algebra learners. In: Coxford A.F. (ed.) (1988). *The ideas of algebras. Yearbook 1988, K-12*. Reston: NCTM.

Knoblich, G., Ohlsson, S. and Raney, G. E. (2001). An eye movement study of insight problem solving. *Memory and Cognition, 29*(7), 1000-1009.

Lakoff, G. and Johnson, M. (1980). *The Metaphors we live by*. Chicago, IL: The University of Chicago Press.

Lakoff, G. and Núñez, R. (2000). *Where Does Mathematics Come From*. New York, NY: Basic Books.

Latour, P. L.(1962). Visual threashold during eyemovements. *Vision Research, 2*, 261-262.

Leung, F.K.S., Graf, K., and Lopez-Real, F. (2006). *Mathematics Education in Different Cultural Traditions- A Comparative Study of East Asia and the West: The 13th ICMI Study*. Springer

Levinas, E. (1989). *The Levinas reader* (Sean Hand, ed.), Oxford, UK: Blackwell.

Lindamood, P., Bell, N., and Lindamood, P. (1997). Sensory-cognitive factors in the controversy over reading instruction. *Journal of Developmental Disorders, 1*, 143-182.

Loftus, G. T., and Macworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance*, 4(4), 565-572.

McCormick, B.H., DeFantim, T.A. & Brown, M.D. (1987). Visualization in scientific computing: Definitions, domain, and recommendations, *Computer Graphics, 21*, 3-13.

Nemirovsky, R., and Monk, S. (2000). "If you look at it the other way…" an exploration into the nature of symbolizing. In: P. Cobb, E. Yackel, and K. McClain (Eds.), *Symbolizing and communicating in mathematics classrooms*, Hillsdale, NJ: Lawrence Erlbaum, 177-221.

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart, and Winston

Radford, L. (2002). The seen, the spoken and the written. A semiotic approach to the problem of objectification of mathematical knowledge. *For the Learning of Mathematics, 22*(2), 14-23.

Radford, L. (2010). The eye as a theoretician: Seeing structures in generalizing activities. *For the Learning of Mathematics, 30*(2), 2-7.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372-422.

Robutti, O. (2006). Motion, Technology, Gesture in Interpreting Graphs. *The International Journal for Technology in Mathematics Education, 13*(3), 117-126.

Scheiter, K., and van Gog, T. (2009). Using eye tracking in applied research to study and stimulate the process of information from multi-representational sources. *Applied Cognitive Psychology, 23*, 1209-1214.

Seitz, J. A. (2000). The Bodily Basis of Thought, New Ideas in Psychology. *International Journal of Innovative Theory in Psychology, 18*(1), 23-40.

Sfard, A. (1991): On the dual nature of mathematical conceptions: reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics, 22*, 1-36.

Sfard, A. (2000). Symbolizing mathematical reality into being: How mathematical discourse and mathematical objects create each other. In P. Cobb,K. E. Yackel, & K. McClain (Eds), *Symbolizing and communicating: perspectives on Mathematical Discourse, Tools, and Instructional Design* (pp. 37-98). Mahwah, NJ: Erlbaum.

Tall, D. (2002). Using Technology to Support an Embodied Approach to Learning Concepts in Mathematics. First *Coloquio do Historia e Tecnologia no Ensino de Matematica at Universidade do Estrado do Rio de Janeiro*, Rio de Janeiro, BR.

Tarmizi, R., and Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology, 80*, 424-436.

van Gog, T., and Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction, 20*, 95-99.

Verschaffel, L., De Corte, E., Pauwels, A. (1992). Solving compare problems: An eye movement test of Lewis and Mayers Cosnsistency Hypothesis. *Journal of Educational Psychology*, *84*(1), 85-94.

Volkman, F. C. (1976). Sacadic suppression: A brief review. In R. A. Monty and J. W. Senders (Eds), *Eye movements and psychological processes.* Hilsdale, NJ: Earlbaum.

Yarbus, A. L. (1967). Eye movements and Vision. New York: Plenum.