CHIARA ANDRÁ, PAULINA LINDSTRÖM, FERDINANDO ARZARELLO, KENNETH HOLMQVIST, ORNELLA ROBUTTI and CRISTINA SABENA

READING MATHEMATICS REPRESENTATIONS: AN EYE-TRACKING STUDY

Received: 5 January 2012; Accepted: 6 August 2013

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 of this study 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 graphical registers.

KEY WORDS: Eye tracking, Mathematical representations, Semiotics, Visual perception

It is well acknowledged that reading a mathematical text requires giving sense to a set of multiple representations of mathematical objects, such as formulas, graphs, drawings and words. However, remedial literacy education on *text reading* has shown that the processes of "decoding" (i.e. recognizing printed words) and "comprehension" (i.e. understanding what words means) are not highly correlated (e.g. Lindamood, Bell & Lindamood, 1997). This urges us to briefly sketch theories that can help us define what we mean for "reading" and what do we mean for "understanding" and the relationships between these two terms. It is also necessary to define the philosophical perspective from which we interpret the theories underlying these two terms. 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 introduction reflects this aim and continues 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 "Overture"

Cognitive psychologists have been interested for decades in text processing in general. Kintsch (1998) argues that text comprehension involves processing at different levels. From the linguistic level, text

Electronic supplementary material The online version of this article (doi:10.1007/s10763-013-9484-y) contains supplementary material, which is available to authorized users.

Published online: 11 December 2013

processing entails decoding symbols to the semantic analysis which determines the meaning of what is read. Forming propositional units according to their syntactic relationships and the coherence of the text also provides us with an understanding of the text. Thus, inserting the text in the (proper) situational model integrates 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. These mental models of the world represent a person's perception. A mental model has the same structure of the situation it represents. However, this model is partial since it represents only certain aspects of what it represents. 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 (verbal and imagery) are activated when one recognizes, manipulates, or thinks about words or things. In reading, when a person combines pictures, mental imagery, and verbal elaboration, he or she understands and learns more effectively than just looking at plain text. These theories from cognitive psychology point out 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 respect to this, the eye tracker as a method may provide useful insights into the way students read mathematical texts. In Radford's (2010) examination of how the eye of a student is *domesticated* by the student's cultural and education, he showed that the student's perceptional view of mathematical objects in particular shapes the student's understanding of the objects as seen within the person's 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 so 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 (e.g. De Freitas & Sinclair, 2012), we see that perceptual routine habits and material interactions constitute conceptual categories. Thus, the focus on the movements of the eye when a student is reading a mathematical text can provide us with insights on his learning processes, such as 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 which allow an individual to see the world. The focus of our study 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. Thus, our purpose as mathematics education researchers is to understand what can and cannot learned from this investigation 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.

The focus of our study is on the reading of formulas, graphs, and plain texts in mathematics. 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. In this study, we are concerned with answering the following questions.

1. Are there differences between formulas and graphs, in terms of the movements of the eye?
2. What do such differences, if any, entail in terms of reading and understanding a mathematical text?
3. 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 sociocultural history of mankind (Radford, 2010). 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). There is no *noesis* without *semiosis* (Duval, 2006); namely, cognitive acts in mathematics such as understanding a concept are not possible without the existence of a relation between the signifier (a sign, a representation) and the signified (the mathematical object) as put forward by Johnson-Laird (1983) in their reference 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. In this sense, taking a semiotic lens for examining mathematics learning is a privileged point of view, given that the highly structured semiotic systems are central to mathematical activity (Leung, Graf & Lopez-Real, 2006).

From our understanding, 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 students takes place when looking at representations in a certain, intentional, and cultural way (Radford, 2010). Mathematical texts are full of complex mathematical representations, such as formulas, graphs, numerical tables, and natural language. Each of these representations must be properly interpreted and coordinated with the mathematical representation for the purpose of complementing, constraining, and supporting the construction of deeper knowledge (Ainsworth, 2006). Moreover, the reading of these representations may be influenced by a variety of factors, such as one's knowledge, the context of the task, the structure of the representation itself, and the mental models activated by students (Johnson-Laird, 1983), how representation is encoded and the role of verbal–imagery components (Paivio, 1971), and how text is decoded and comprehended at various levels (Kintsch, 1998).

Cognitive load theory is concerned with the ways instructional materials can enhance or impede learning. Tarmizi & Sweller (1988), in the learning of geometry, found that students' attention split towards an integrated representation enhanced learning as 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 & Sweller (1991) concluded that the integration of multiple sources of information was beneficial when representations are unintelligible without being integrated. This is in accordance with the aforementioned semiotic approach where 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 reveals itself to the learner* and how two cognitive processes turn out to be important when focusing on the visual experience of a subject to be learned. The first one is concerned with the process of attending the representation in order to access its meaning; the second one is concerned with the shift from one representation to another representation while focusing on the same mathematical object (Duval, 2006). Namely, how a mathematical representation is seen as a cultural product, with its own rules, letting the representation itself 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 our study within a cultural semiotic perspective since mathematical objects are a product of our culture. However, 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) mathematical representations act as a cultural product reveals itself to the learner, the research methodology used in this study was the eye tracker. The eye tracker allows us to track the movements of the eye while a participant is performing a task. The eye-tracker methodology consists of a camera, which is filming the movements of the eye. Studies in neurosciences have shown that human cognition is embodied: it takes place not only in the areas of the brain that are delegated to high mental processes, but also in those that are proximal to the ones related to sensorial experiences, body movements, and actions. Cognition is strictly grounded in the human body and in its location in space and time (e.g. Lakoff & Núñez, 2000). Thus, eye movement is part of sensorial experience. 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 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 what is to be learned and as a means to express mathematical problems (Duval, 2006).

But plain text is also the mode of expression of other activities which lie at the core of mathematics, such as exploring, conjecturing, problem solving, arguing, and proving. Kintsch (1998) argued 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 and leaving aside an examination of students' problem solving processes. 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 have also emphasized that mathematical symbols represent both “processes” and “objects” and that this dual nature of mathematical symbols is difficult to manage by students (Sfard, 1991). To highlight the amalgam of the dual natures of process and object within symbols, Gray & Tall (1994) spoke of a “procept.” They defined 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.

The theoretical framework provided by Kintsch (1998) on textual comprehension at different processing levels applies to comprehension of mathematical formulas as well: (1) decoding symbols and recognizing their role within the syntax of the formula. Since 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”). Formulas often have a nonlinear structure (e.g. the structure of the formula $3x+2[x+5y]$) which do 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 ($3x+2x+5y$); (2) perceiving the global sense of a formula out of any single symbol; and (3) situating it with a proper model. A formula is supposed to be read according to the rules given by its semiotic register (Duval, 2006). In the passage from text to formula, a *conversion* from one semiotic register to another semiotic register 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 information in the plain text and to do it symbolically and according to formal rules. In the opposite direction, the conversion from formula to plain text requires unpacking of such symbolic information, in order to give a narrative and discursive understanding of it. Specifically, in our study, we looked at formula–plain text conversions in both directions. As Radford stressed in several works (e.g. Radford, 2002, 2010), a crucial problem for learners is to understand 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 merely substitutes of nouns but are the means by which objects of discourse function (Radford, 2002).

These theoretical premises not only lead us to claim that the conversion from text to formula and the conversion from formula to text are a central cognitive act in the learning process of those areas of mathematics where symbols play a significant role, such as algebra, but also urges us to investigate 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. For example, we just quoted 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 *into something else* within the same symbolic register. These "manipulative" and demanding features of formulas do not seem to belong to graphs in mathematics where students are often expected to interpret them. Graphs can represent other kinds of modeling the world (Johnson-Laird, 1983) in mathematical terms. While in formulas the verbal component is dominant, the imagery component is dominant in graphs (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 are not subsequent to unpackaging (like formulas). In addition, graphs have an iconic component that is absent in formulas. As a graph changes, its size also changes. Learning how to interpret a graph, a student has to integrate all the elements of the graph into a comprehensible whole (Robutti, 2006). If students are not taught how to perform such integration, students may focus only on certain parts of the graph without acknowledging that there may be other useful and necessary information elsewhere in the graphic representation. In order to make sense of a graph, a suitably strategy can be employed to understand the narrative by linking the shape of the graph to the phenomenon it is modeling. Nemirovsky & Monk (2000) presented an interesting example of modeling motion. 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. Similar to formulas, reading a graph requires a subject to locate and decode the position of relevant information (e.g. Carpenter & Shah, 1998). A graph usually condenses a lot of information, which is not always explicit. 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 by 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.

Experimental Design and Methodology

Based on our previous review of literature, 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 selected three conversions: from formulas to plain texts, from plain texts to formulas, and from graphs to plain texts. Also, other conversions are crucial in mathematics. These additional conversions are from text to graph and from formula to graph, or vice versa. 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. These stimuli were given to the students in random order for the purpose of avoiding fatigue among the students. These three types were (1) formula–text (ft type, $N = 15$ items) where the input is a formula with four text alternatives, (2) graph–text (gt type, $N = 12$ items) where the input is a graph with four text alternatives, and (3) text–formula (tf type, $N = 16$ items) where the input is a text with four formula alternatives. The students' task was to determine which of the four alternatives is an accurate representation of the input or corresponds to the input according to the rule or feature of the representation. The correct answer was always presented among the alternatives.

We decided not to 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 and the writing process when students are constructing an answer.

The focus of our study 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 which allows the identification of the parts of the stimulus that contain relevant *areas of interest* (Knoblich, Ohlsson & Raney, 2001; Verschaffel, De Corte & Pauwels, 1992; Hegarty, Mayer & Green, 1992). By using an area of interest with a semantic mathematical content, we are able to identify something specific about how information in that area is processed. For each stimulus, we consider five areas of interest. One area of interest is the input while each of the four answers represents an area of interest. The example in Fig. 1 shows a stimulus that requires a conversion from a formula to a plain text. The input contains a formula, and the four alternatives contain plain text. The plain text may contain some symbolic expression, but the overarching cognitive process is to pass from the formula to the 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 (i.e. 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. These levels are the *macro*level (e.g. the five big areas: input and four alternatives), the *meso*level (e.g. within each of the five areas), or the *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 Andrà et al., (2009). Since the purpose of our study is to understand how the representation *as a whole* reveals itself to the learning subject and how the subject makes sense of the representation by selecting one out of four alternatives rather than by reading or navigating with the eye a certain area of the stimulus, we operated at the macrolevel which focused on the interest area of the whole input/alternative. We use areas of interest to examine "overall behavior" since we are not looking at the detailed behavior of every single transition. Even though performing such a detailed analysis of every

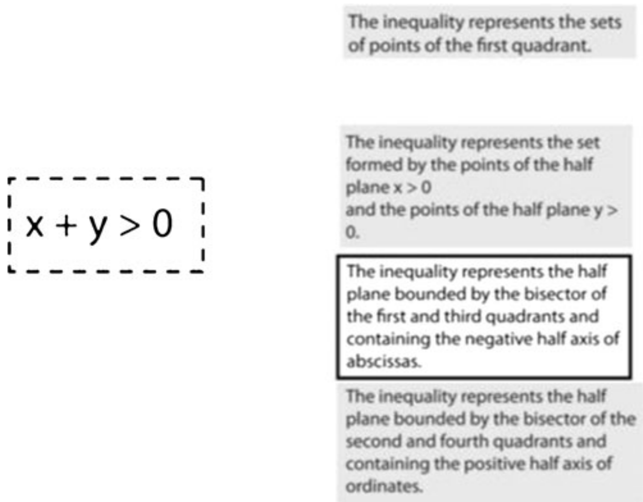

Figure 1. Five areas of interest. The *dotted box* corresponds to the input. The *black box* corresponds to the correct alternative, and the *gray boxes* refer to the incorrect alternatives

single transition is possible, such a detailed analysis was not within the scope of our study. 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 et al., 2001) have done. Such studies have shown in word problems that certain areas are useful to attend to in order to solve them (Hegarty et al., 1992; Verschaffel et al., 1992).

All of the 43 stimuli which were used in this study are located on the Electronic Supplementary Material Website. 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 over another. Part of this process consists of navigating the area of interest with the eyes according to the semiotic register. Such navigation during eye movement can differ. Since navigation is not merely 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 as follows: Given the diversity of the stimuli in both complexity and difficulty, is it possible to detect such differences in attending different semiotic registers? In order to answer this question, we used eye-tracking methodology which allowed us to consider variables that may vary according to the different types of stimuli.

Forty-six Swedish university students were recruited based on their different knowledge backgrounds in mathematics. Twenty-four students had not studied mathematics at the university level. Twenty-two students had studied 1 year of mathematics in the engineering department. The differences in performance between students who had studied and who had not studied mathematics at the university level are shown in Andrà et al., (2009). As stated earlier, in this study, we concentrated only on the differences between representation types of the stimuli. This is also in accordance with the question presented at the end of the previous paragraph. Our aim was 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 did not classify stimuli according to level of difficulty and/or complexity. As well, a literature review on these issues is not in line with the interest of our research and thus was not presented. We, instead, relied 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 and provide a better understanding of how the students deal with semiotic representations.

Now, we present and discuss the variables, from the eye-tracker methodology, used 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, the fixation duration was calculated as the average of all fixations in a given area. If we have two fixations of, respectively, 200 and 300 ms, then the average fixation duration was calculated as 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 average fixation durations for formulas are the longest, followed by graphs, and the shortest for text. 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 figurative/symbolic and textual components which 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 et al., 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. We also predict that the most number of fixations will be on 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 two 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 graph as shown by Carpenter & Shah (1998). 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 indicate that the subject identifies the relevant information in the formula more quickly than that in the graph, while a longer dwell time on the text input area may just confirm that text requires some time to be read. A comparison among these three kinds of stimuli may reveal how long each one is attended, beyond its structure and its features.

The three measures previously 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 which 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 (Duval, 2006). Since this is an exploratory study on a macrolevel with respect to the area of interest, we further 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 and focuses 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 behaviors 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. Pair-wise comparison can be traced back to Kintsch's (1998) text processing at the microlevel, which searches for and analyzes the coherence relations among different areas of the stimulus. Overview looking can refer to the macrostructure which entails also the recognition of global topics and their interrelationships (Kintsch, 1998). According to Kintsch, the microstructures and macrostructures together form the text base 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 movements 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, a student is performing a pair-wise comparison between the input (I) and the first alternative (A). While in the second part of the sequence, the student is navigating all five areas (i.e. the input and its A, B, C, and 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 stimulus. 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 is 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, and GT12; the easiest gt-type stimuli were GT02, GT16, and GT18. If one looks closely at them (on the Electronic Supplementary Material Website), one concludes that there is not a strikingly common feature, in terms of eye movements. However, we found that we could 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 were not “easier to be read” but easier to be dealt with in mathematical and didactical terms. They required commenting on the trend of the graph. Conversely, difficult gt-type stimuli required students to make some computations or to have some knowledge of calculus. This may help to confirm our starting hypothesis in the subsequent analysis which was that 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 were 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, and FT16) required some reasoning, while the easiest ones (FT02 and FT04) required commenting on some features of the formula given in the input.

Bar charts in Fig. 2 showed the average and had 95 % confidence intervals for dwell time, fixation duration, and number of fixations. Since, in almost all the cases, they do not overlap, we inferred that differences between types were almost always significant. In Fig. 2 (left), the longest total dwell time for the graph input is 11.2 s on average. This is followed by the text (7.9 s) and then the formula (7 s). Since participants were given 40 s maximum for attending the whole stimulus, the total dwell time on the input area is not less than 6 s and not more than 12 s, which corresponds to 15 and 30 %, respectively, of the total time per stimulus.

From Fig. 2 (middle), the longest fixation durations are for the formula input, and the shortest fixation durations are for both the graph and text inputs. The fixation duration *on average* is not shorter than 190 ms and is not longer than 250 ms. The range is quite narrow and expressed in milliseconds! Looking at the width of confidence interval, the variability with respect to this variable is low. The width of the confidence interval is 2–4 ms. All of the stimuli are of the same type and have very similar fixation durations. Looking at Fig. 2 (right), there is the greatest number of fixations when the input is a graph, followed by the text input and by the formula input.

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 alternatives. Figure 3 (left) shows that there are some differences between the cases in which a graph is present with respect to formulas regarding the pair-wise comparisons. The gt-type stimuli have lower pair-wise comparisons in percentage. In Fig. 3 (right), we see that there is more overview looking for the ft type as compared to the gt type and the tf type. Looking at the percentages range, the pair-wise comparisons and

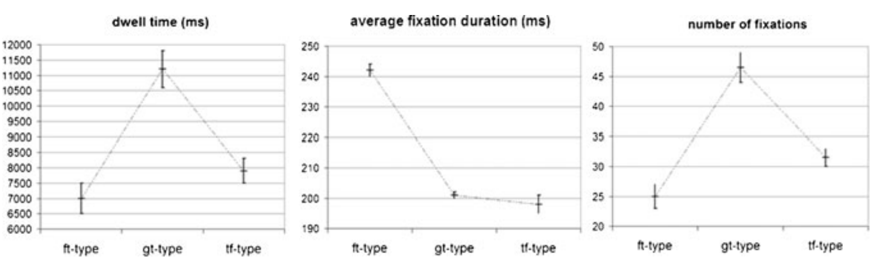

Figure 2. Different representation types: (*left*) dwell time, (*middle*) number of fixations, and (*right*) average fixation duration. Ninety-five percent confidence intervals are reported and show significant differences between the estimated values

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 macrolevel. 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 subparts and considers transitions between the subparts, which can be as fine as the pixel definition of the screen. This is not surprising given the kind of areas the students were presented. Moreover, with reference to the data from dwell time above, we know that 15–30 % is the proportion of time the students looked at the input. Hence, the remaining 45–60 % is the proportion of time in which participants looked at the four alternatives.

To sum up, we have observed that, for gt-type stimuli, the dwell time for this input is the longest with the highest number of fixations, but the average duration of each fixation 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 have lower percentages; hence, we can conclude 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 dwell time for this input is the shortest with the lowest number of fixations. The average duration of each fixation is the highest. 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 for a relatively longer time but with few times of observation for

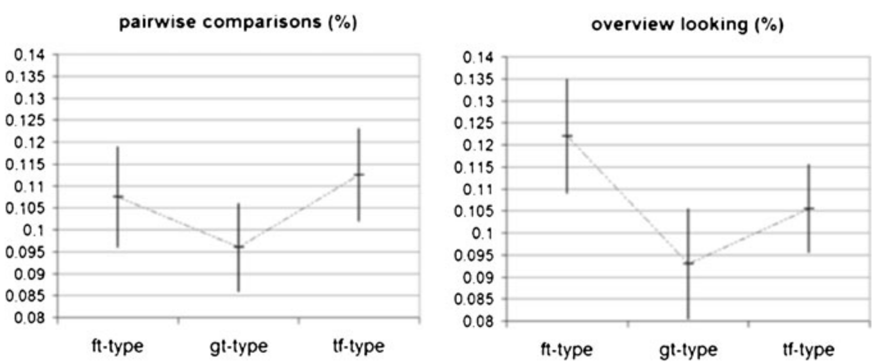

Figure 3. Percentages and confidence intervals of (*left*) pair-wise comparisons and (*right*) overview looking for the three kinds of stimuli

the purpose of navigating 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 student needs fewer but longer fixations since she/he focuses on the specific relevant elements (e.g. an exponent, a sign, or a coefficient). A graph, on the other hand, is usually perceived in a more holistic and condensed (synthetic) manner, which requires more but shorter fixations by the student subjects.

As regards tf-type, some values are close to ft-type, others are different. The dwell time is short but not as short as the ft-type case. The fixation duration of the tf-type is very low, and its number of fixations is in between gt-type 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 the ways one perceives formulas and graphs. In a single sentence, one can say that, on average, in reading *formulas*, students have a *smaller number of longer fixations* than those in *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. The dwell time, in fact, is higher for graphs, telling us that the big number of short fixations on a graph results in an overall longer dwell time 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 representing that a formula condenses its information in a shorter but more hidden inscription. Both verbal and imagery components (Paivio, 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, the person will in turn have difficulty 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. As also Mason (2008) stressed, the learners' primary concern is achieving satisfactory answers to tasks. The students will usually act as if they know a theorem/ property or a technique/algorithm, but there is no evidence that they are aware of it at more than a functional level. In a sense, while Radford (2002) signals that it is not enough to stand in front of a mathematical object and look, it is, however, necessary to see beyond the crude perception to find something that was previously unnoticed. For example, in algebra, it is necessary to understand that letters are not mere substitutes for nouns and require the use of a different kind of designation for the objects of discourse. Mason (2008) argued that attention cannot be reduced to its behavioral manifestation, but it is *observation* itself, as well as awareness which cannot be reduced to the individual's ability to speak about what she/he is doing, but may reside in our body as we may initiate an action that our consciousness is able to build a narrative on which only later in time. Automatic strategies and speedy elaborations can be seen on the side of the didactic contract (Brousseau, 1997). Students believe that they are actually learning when they complete the task as best they can (Mason, 2008). The large majority of eye movements in our study seems to confirm that the procedural nature of the way the learners come to deal with formulas often impedes the meaning-making processes that are crucial for making sense of symbols and their syntactic organization to understand a formula.

However, graphs turned out to be the easiest, and pair-wise comparisons and 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 a matter of time but 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). However, Roth (2003), in analyzing how graphs are collectively interpreted and made meaningful, stresses the importance of the *context* in which graphs are produced and looked at even when the individuals who enter in relationship with such graphs belong to a community of specialized field of expertise. Roth stated that “graphs as objects do not exist as independent entities, but are complex networks that integrate entities and processes” (p. 305). According to Roth, graphs do not speak by themselves (in fact, graphs do not invite to automaticity in terms of how to read them), and their interpretation can rather turn out to be a projection of what the scientist who produced it is looking for. Moreover, the competence of the scientific reader of a graph strongly depends on the reader's proximity with the aforementioned network of entities. The interpretation of a graph is meant as an historical process in which a system of signifiers and processes stabilizes itself. Thus, graphs contribute to and create a shared interactional space in a continuously transforming social network. Our study confirms Roth's warning that teachers should be aware that graphs are not self-evident, and it is necessary that the learners develop the ability not only to look at graphs but also to interpret them as well as to share and to discuss the details that are relevant for each graph. Moreover, we have shown that graphs do not invite the reader to compare them with their textual counterparts. These facts point out the need of educating the eye—as outlined by Radford (2010)—and this process of education can take place only within a social environment where the learners are engaged in meaningful mathematical activities.

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 led us to the conclusion 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-type 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, the textual input is attended by the students in a specific way. To use Mason's (2008) words, overview looking can be meant as a form of attending that consists in gazing and waiting for things to come to mind (and in fact it is higher in ft-type stimuli, since the formula in the input requires a certain degree of confidence with the rules regarding its reading), whilst pair-wise comparisons can be traced back to discerning details, a structure of attention that allows the learner to become familiar with the worthwhile features to look out for: tf-type stimuli seem to invite the reader to discern details of both the textual and the symbolic representation of the mathematical object, in order to grasp the correspondences between the two.

These summarized results allow us to answer affirmatively our research question that 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). The eye-tracker methodology provided us with suitable tools 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 also mostly as culturally and socially shaped sense which allows us attending mathematical representations in a certain structured way. The foundational basis of our study is phenomenological in nature (Mason, 2008). Mason understands observation and attention as synonymous and the significance of attention. This is not only *what* is attended to, but this is also about the nature (attention can vary in multiplicity, focus, locus, and sharpness) and the structure (holding wholes, discerning details, recognizing relationships, and perceiving properties) of the attention. Learning is an education of awareness which is closely related to attention/observation. Since attention is "where we are," Mason argues that attention not only determines but it is also the mental world we occupy. This world of the learner is essentially a world that focused an attention to the tasks. However, it needs to reach out to the underlying awareness. According to Mason (2008), teaching is about directing the learner's attention and about being aware of what learners are not yet aware of, and finding ways to prompt them to become aware. Radford (2010) shared a consonant perspective and pointed out the need of educating the eye, hence the student's attention. In order to trigger and foster the learner's process of objectification, a student must be aware of the mathematical activity in which the student is engaged. The teacher, as Mason (2008) argued, can direct students' attention, especially when they are being mathematical with and in front of their learners. In this perspective, Arzarello et al., (2011) adopted a compliant phenomenological lens to analyze two teachers' classroom activities regarding the graphical antiderivative and suggested to take Schoenfeld's (2010) ROG—Resources, Orientation, and Goals—framework to see that teaching begins with the teacher's orientations, which shape the prioritization of the goals to be established and, thus, the relevant resources (one for all, knowledge) to be activated. Regarding the education of the eye, we would like to end this work acknowledging the complexity of this process of education/“domestication” (Radford, 2002) and call for further researches, from more than one field of knowledge, in this direction.

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. & 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.

Arzarello, F., Ascari, M., Thomas, M. & Yoon, C. (2011). Teaching practice: a comparison of two teachers’ decision making in the mathematics classroom. In: B. Ubuz (Ed.), *Proceedings of the 35th conference of the international group for the psychology of mathematics education* (Vol. 2, pp. 65–72). Ankara, Turkey: PME.

Brousseau, G. (1997). *Theory of didactical situations in mathematics*. Dordrecht: Kluver.

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

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

Chandler, P. & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*(4), 293–332.

De Freitas, E. & 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.

Gray, E. & Tall, D. (1994). Duality, ambiguity, and flexibility: a “proceptual” view of simple arithmetic. *Journal for Research in Mathematics Education, 25*(2), 116–140.

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.

Henderson, J. M., Weeks, P. A., Jr. & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene 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. & Sabena C. (2011). A method for quantifying focussed versus overview behaviour in AOI sequences. *Behavior Research Method, 43*(4), 987–998.

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

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

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

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

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

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

Lakoff, G. & Núñez, R. (2000). *Where does mathematics come from*. New York: Basic Books.

Latour, P. L. (1962). Visual threshold during eye movements. *Vision Research, 2*, 261–262.

Leung, F. K. S., Graf, K. & 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). In: Sean Hand (ed), *The Levinas reader*, Oxford, UK: Blackwell.

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

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

Mason, J. (2008). Being mathematical with and in front of learners: attention, awareness, and attitude as sources of differences between teacher educators, teachers and learners. In T. Wood & B. Jaworski (Eds.), *International handbook of mathematics teacher education: the mathematics teacher educator as a developing professional* (Vol. 4, pp. 31–56). Rotterdam: Sense publishers.

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

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.

Roth, W. M. (2003). *Toward an anthropology of graphing: semiotic and activity-theoretic perspective*. Dordrecht: Kluwer.

Scheiter, K. & 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.

Schoenfeld, A. H. (2010). How we think: a theory of goal-oriented decision making and its educational applications. Routledge, New York.

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: 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. & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology, 80*, 424–436.

van Gog, T. & 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 consistency hypothesis. *Journal of Educational Psychology, 84*(1), 85–94.

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

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

Chiara Andrá, Ferdinando Arzarello and Ornella Robutti Cristina Sabena

Dipartimento di Matematica
Università di Torino
Via Carlo Alberto, 10, 10123 Torino, Italy
E-mail: chiara.andra@gmail.com

Paulina Lindström

Lund University Cognitive Science
Kungshuset Lundagård, 222 22 Lund, Sweden

Kenneth Holmqvist

Humanistlaboratoriet, SOL Centrum
Lunds Universitet
Helgonabacken 12, 223 62 Lund, Sweden