The role of semiotic resources when reading and solving mathematics tasks

ANNELI DYRVOLD

One part of being proficient in mathematics is to be able to read and solve mathematics tasks where mathematics is represented using different semiotic resources (i.e. natural language, mathematical notation, and different types of images). In the current study, statistical methods are used to investigate the potential meaning that the presence and co-occurrences of semiotic resources have for how demanding a mathematical task is to read and solve. The results reveal that the number of different semiotic resources in a mathematical task is not related to difficulty, but that difficulty is related to the particular combinations of semiotic resources where pictorial images are one of the resources. The results also indicate that the difficulty related to these semiotic characteristics is not related to an unnecessary reading demand.

The concept *representation* is well established in mathematics education research and it is used both for internal and external representations (e.g. Goldin, 1998; Hiebert & Carpenter, 1992). The current article is about the external representations but the notion of *semiotic resource/s* is used instead to stress that the study regards differences between four semiotic resources that can be present in the same representation. The four semiotic resources are natural language (words and letters), mathematical notation (symbols that are not Latin letters), and two types of images: schematic images like tables and graphs, and pictorial images that are naturalistic and detailed. The four semiotic resources are defined under Method. The word representation can refer to a textual element with multiple semiotic resources (e.g. a circle labelled with the text ”$A = r^2\pi$”), and thus the word ”representation” could obscure what is exactly referred to. The terms ”representation” and ”semiotic resource” often function synonymously, but to maintain consistency in the text, the term ”semiotic resource” is used. What the task text looks like regarding presence and

Anneli Dyrvold
Umeå universitet

Dyrvold, A. (2016). The role of semiotic resources when reading and solving mathematics tasks. *Nordic Studies in Mathematics Education*, 21 (3), 51–72.

co-occurrence of different semiotic resources is in this study referred to as the tasks' *semiotic characteristics*.

Engaging in mathematics inevitably means dealing with different semiotic resources and the use of those resources is, thus, one part of how we communicate in mathematics. The ability to read and use different semiotic resources are also mentioned as part of various mathematical competencies (see e.g. NCTM, 2000; Niss, 2003) and to be fluent in the use of several semiotic resources can therefore be perceived as part of being mathematically proficient. Research has pointed to the central role that the multisemiotic language has within mathematics (Lemke, 2003; O'Halloran, 2005; Pimm, 1995), and it has even been proposed that it is in the interactions between several semiotic resources and what is being referred by these semiotic resources that mathematics is created (Pimm, 1995). In addition, the grammar of the mathematical notation system enables problem solving in mathematics that would not be possible with other semiotic resources (O'Halloran, 2005). For a mathematics text with several semiotic resources, it is often evident that the different resources are needed for different purposes. Lemke argues that it is impossible to construct the same meaning with two different semiotic resources and that the use of multiple semiotic resources might be needed for certain meanings to be stated (Lemke, 1998). If such a view is adopted, then the way the semiotic resources are used in mathematics is seen not just as a way to express mathematics but as an intrinsic part of the mathematics itself.

With the use of functional linguistics, O'Halloran elucidates several aspects of the complexity with which different semiotic resources interact to express meaning in mathematics discourse and how the meanings made in interaction with several semiotic resources are multiplicative (see e.g. O'Halloran, 2005, 2007, 2008). In this context, the concept of *multiplicative* means that what can be expressed with the semiotic resources together is more than the sum of what is possible with each of the resources alone. In the analysis of a mathematical task, O'Halloran (2008) identifies seven intersemiotic mechanisms which enable meanings to be made, that would not be possible with natural language alone.

Based on this previous work, it is evident that the presence of several semiotic resources is important for the meaning being conveyed in a multisemiotic mathematics task. This means that the ability to read and solve *multisemiotic* tasks is one important part of a communicative competence in mathematics. However, not much is known about what this competence comprises and therefore the current study is concerned with difficulties related to the reading and solving of tasks with different semiotic characteristics. A restriction to presence and co-occurrences of different semiotic resources is considered an appropriate first step in investigations about abilities needed when reading and solving multisemiotic mathematics tasks. Results that have to do with the tasks' multisemiotic characteristics in relation to how difficult the task is to read and solve are of interest from an educational perspective since it gives implications about what students might struggle with. Tasks from the Swedish national tests in mathematics for grade 9 (hereafter referred to as SweNT) and PISA mathematics tasks (Programme for international student assessment) are chosen as data since it is two large samples of tasks for which solution frequencies for many students are available.

Background

Duval (2006) argues for the essential role that semiotic resources have had for the historical development of mathematical thought and that they still have for all work with mathematical objects. Semiotic resources are essential because working on and thinking about mathematics inevitably means engaging in transformations between different semiotic resources, and such transformations can be sources for incomprehension (Duval, 2006). Contributing to the complexity of mathematics is that mathematical objects do not exist as physical objects (Duval, 2006; Sfard, 2008). Mathematical objects are abstract knowledge objects, for example concepts or procedures. Duval's analysis of comprehension in mathematics reveals several ways in which the abstractness of mathematical objects can be a source for incomprehension. The crucial part is that students have to handle representations of mathematical objects without access to the mathematical object, and they must rely only on the semiotic representations. For example, when transformations are made between different semiotic resources (such as an algebraic expression of a function or a graph), a strict understanding of the object as the representation can lead to a conflict when a transformation of the object to another semiotic resource is required. Two representations of the same object can then be misunderstood as two separate mathematical objects (Duval, 2006). Therefore, students have to use semiotic resources in every mathematical activity, but at the same time they must not confuse the representation with the mathematical object.

Because the language of mathematics has developed as a multisemiotic language (O'Halloran, 2005; Radford & Puig, 2007), it is important to understand the potential difficulties related to the use of, and transformations between, semiotic resources. Mathematical objects and how they are represented are also a focus in Sfard's research (Sfard, 2008). She argues that learning mathematics means developing a discourse and working with discursive objects. The discursive object is a personal construct made up of all the realizations the individual implements as belonging to the discursive object (Sfard, 2008). With a view such as Sfard's on what it means to understand mathematics and discursive objects, the interpretation of mathematics texts (as a mathematics discourse) means not just interpreting the separate representations in the text, but also understanding which discursive objects the realisations are a part of and how they are related. Despite different theoretical perspectives, both Sfard's and Duval's argumentation contributes to an understanding of the semiotic characteristic of mathematics texts as essential for learning and communicating in mathematics.

Reading mathematics texts often means reading images like diagrams together with natural language. Analyses of how students read mathematics texts reveal a significantly greater use of inferences and high-level strategies when diagrams are read than when reading natural language (Cromley, Snyder-Hogan & Luciw-Dubas, 2010). The fact that the interpretation of images can be demanding has also been demonstrated through an analysis of students' solutions to mathematical tasks, where many errors could be attributed to an inability to correctly decode images – especially conventions used to visualise movements and reflections – or to discriminate between different images (Lowrie, Diezmann & Logan, 2011). However, images do not need to have complex features to interpret for them to be considered difficult. Statistical analyses of the time used and the accuracy of performance on mathematics tasks with different types of images reveals that images with redundant information also have the potential to negatively influence the ability to solve arithmetic problems (Berends & van Lieshout, 2009).

It might be tempting to explain away the relations between particular semiotic resources in mathematics tasks and task difficulty by arguing that the relation stems from a more frequent use of some semiotic resources in more difficult areas of mathematics. Such a relation is possible, but it is not the only explanation for the difficulty in relation to the semiotic resources that are present in the task text. Earlier research has taken different perspectives to show that the existence of both different types of images, and mathematical notation in mathematics tasks is connected to how demanding the tasks are to read, comprehend, and solve. For example, tasks testing numerosity are more demanding when Arabic numerals are used than when images are used to represent numerosity (Rousselle & Noel, 2007). The presence of images can also be related to task difficulty depending on the type of image. Statistical relations between the semiotic resources that are present in mathematical tasks and the performance on the tasks support the view that tasks with informational images (as opposed to decorative images) are more demanding compared to tasks with similar mathematical content but without such images (Gagatsis & Elia, 2004). A comparison between students' achievement on 32 tasks where the item response options are given as mathematical notation (numerals and other mathematical notation symbols) or as schematic images showed that students scored significantly higher on the test with mathematical notation than the test with schematic images (Lin, Wilson & Cheng, 2013). The existence of mathematical notation in texts has also proven to be demanding in a study comparing comprehension of the same mathematical content in text with or without mathematical symbols (Österholm, 2006). In conclusion, the presence of different semiotic resources in tasks seems to be connected to how demanding the tasks are for the students.

Analyses of students' errors in the translation between different semiotic resources reveal the following three sources for the errors: students have weak reading comprehension; they lack knowledge of symbols, words, and algebraic representations, and they have deficiencies in their ability to use and organize information from different parts of a task when the information is given using different semiotic resources (Duru & Koklu, 2011). These three types of sources for errors might also be part of the explanation for the result of a study by Mundy and Gilmore (2009) that found a positive correlation between high achievement on mathematical tests and the ability to translate between semiotic resources. Another possible way of interpreting this result is that the ability to move fluently between semiotic resources is one part of mathematical competence.

Students have different strategies for reading texts with several semiotic resources, and some of these strategies might lead to misunderstandings. For example, diagrams might be skimmed or skipped (Cromley et al., 2010; Gagatsis & Elia, 2004), in some cases in favour of some intuitive or everyday understanding of the content in a task (Lowrie et al., 2011). There are also results showing that students sometimes choose to read from one type of representation because they experience difficulties in using disparate types of representations and translating between them (Ainsworth, Bibby & Wood, 2002). Another reading strategy leading to misunderstandings is to favour the natural language in the text and to pay less attention to different types of images (Elia, Gagatsis & Demetriou, 2007; Gagatsis & Elia, 2004). Analyses of how students read mathematics tasks while solving them reveal other inefficient strategies, including skipping or replacing difficult words or mathematical notation with familiar words and using everyday understanding to solve tasks instead of using the mathematical content presented in the text. These are strategies that lead to a lack of deeper levels of reading comprehension (Adams & Lowery, 2007). All of these strategies are signs of deficiencies in mathematical proficiency because working with several semiotic resources is such an important part of mathematics.

The opposite of inefficient reading strategies would be some kind of reading proficiency, apt for mathematics. This aspect of reading mathematics, as opposed to reading natural language, has been elucidated from several perspectives (see e.g. Adams, 2003; Bergqvist & Österholm, 2010), and it has been shown that language ability (in the form of vocabulary and listening comprehension) predicts gains in some areas of mathematics but not, for example, in arithmetic or algebra. This suggests that language ability influences how students make meaning in mathematics, but it is less influential for dealing with complex arithmetical procedures (Vukovic & Lesaux, 2013). Another way of interpreting this result is that language ability (as used to read natural language) is only partly useful when dealing with mathematics text with different types of images and mathematical notation and that another kind of reading ability is also needed while reading mathematics.

In the current study mathematical tasks' semiotic characteristics is investigated in relation to how demanding mathematics tasks is to read and solve. Therefore, a measure of the non mathematics specific *demand on reading ability* (DRA) in mathematics tasks is used (see Method). The use of this measure together with measures for task difficulty enables interpretations regarding a mathematics-specific difficulty when an unnecessary demand on reading ability is excluded (see also Dyrvold, Bergqvist & Österholm, 2015).

In essence, focusing on the way the presence of semiotic resources in mathematics tasks can be demanding is justified both theoretically and empirically. Theoretically, the elusive nature of mathematical objects implies that students need to recognize the same object represented by various semiotic resources and that the only way to access the object is through semiotics because it does not exist as a physical object (e.g. Duval, 2006; Sfard, 2008). In addition, empirical results illuminate several aspects of difficulty in relation to the presence of different semiotic resources in mathematics tasks, and furthermore, there are some contradictions in the results, thus the difficulty of a mathematics task in relation to its semiotic characteristics appears to be more nuanced than might originally be expected.

Purpose and research question

The purpose of this study is to enhance our knowledge about if and how the presence of different semiotic resources in mathematics tasks is related to how demanding a task is to read and solve.

> Is task difficulty or the demand on reading ability in any way related to i) the presence of different semiotic resources, ii) the co-occurrence of different semiotic resources, or iii) the number of different semiotic resources?

Method

Data

Mathematics tasks in Swedish from two different tests are analysed: the annual Swedish national test in mathematics (SweNT) from the years 2004–2013 (364 tasks in total) for 15-year-old students (school year 9) and the PISA tests from 2003 and 2012. The years 2003 and 2012 are chosen because mathematics was the subject in focus for the PISA test in those years. The PISA tests from 2003 and 2012 together contain 133 different mathematics tasks. Two different samples of PISA tasks are used. In the analysis of task difficulty, all 133 different PISA tasks are used, and in the analysis of the tasks' DRA, 105 tasks are used in the analysis. In one of the analyses, 27 tasks are excluded because they have a negative loading value for their DRA, and one is excluded because of too many missing values. The measures are based on results from around 1,500 students on each PISA task and around 2,000 students on each SweNT task. Two different tests are analysed since this gives the possibility to evaluate the reliability of a statistical relation (a reliable relation is expected to be significant in both samples). Besides that the two samples are used since they meet different requirements, for the PISA test the DRA can be analysed, and the SweNT is useful since it is a large sample of tasks.

The PISA test and the SweNT test are analysed separately because they are composed differently. The PISA tests are composed of a sample of tasks with broadly the same design, but the SweNT is composed of four different parts, often tested at different occasions. One part is oral, one part requires only short answers, one part demands a written solution, and one part consists of only one extensive problem. The oral part of the test is excluded from the data used in the current study because it to a large extent measures something different from all other tasks, namely oral performance.

Four semiotic resources

In order to fulfil the purpose of the study, all textual elements (the constituents that make up the text) are categorized as one of four semiotic resources depending on *form*. In some cases, a symbol such as "!" can be either of two semiotic resources depending on *context*.

Natural language is defined as language in the form of sentences, phrases, single words, or even single letters. Specific for natural language is that the printed text responds to a verbatim representation (e.g. Kintsch, 1998), and the letters and syllables to specific speech sounds. Units such as abbreviations are also categorised as natural language, for example, ***m***2 is composed of both natural language "***m***" and mathematical notation "2".

Mathematical notation is defined as symbols that are used following special conventions in mathematics. Four categories of symbols are seen as mathematical notation: logograms (e.g. π and $\div$), pictograms (e.g. $\|$ and $\angle$), letters (e.g. AB, a, and β), and punctuation marks (e.g. ! and]) (adopted from Pimm, 1987). One exception from these four types is made in the definition of mathematical notation in the current study; Latin letters used as mathematical notation are categorised as natural language based on the assumption that reading Latin letters in natural language is similar to reading Latin letters in mathematical expressions because the letters correspond to the same speech sound in both semiotic resources.

Images are textual elements that have qualities of likeness, resemblance and similitude in common (Mitchell, 1986). Images have frequently been categorised in earlier research as pictorial images or as schematic images (Hegarty & Kozhevnikov, 1999; Martiniello, 2009; van Garderen, 2006), the categories used in the current study. Pictorial images are images depicting objects or details of objects. *Pictorial images* use the visual medium to image the object/s in a naturalistic way, where the likeness of the object to reality is prioritized and where relations within or between parts of the image are not emphasized. *Schematic images* are images that visualize the way parts (e.g. objects, people, events, or data) are related. The relations can occur between or within objects. For relations within objects, what characterizes the image as schematic is that it is visualised in a way that distinguishes the relations within it. This visualisation can be done by stripping the object of irrelevant features, for example, to visualise a tent as a pyramid without zippers or guy lines. Diagrams are schematic because they visualise relations with means that do not exist as physical objects (e.g. a graph representing the acceleration of a car). Images can sometimes be pictorial and schematic at the same time if they take both roles and are then categorized as both types (e.g. exploded view drawings).

Measures

In order to analyse all different aspects of the presence and co-occurrences of semiotic resources in the tasks, twelve semiotic characteristics are defined (table 1). The first characteristic, natural language present in the task (N), is not used in the analysis because all tasks except one contain natural language. An ”o” in the variable name means *only*, for example, ”oNM” describe tasks with only the combination natural language and mathematical notation. The first four variables (N, M, P, S) differ from the rest because the variable only says something about the presence of one semiotic resource. For example tasks in the group ”M” are all tasks that have mathematical notation present, but any other semiotic resource is present can vary in that group.

Table 1. *Variable names for each semiotic characteristic and explanations.*

Variable name	Semiotic resources present in the task text
N	Natural language
M	Mathematical notation
P	Pictorial images
S	Schematic images
oN	Only Natural language
oNM	Only Natural language and Mathematical notation
oNP	Only Natural language and Pictorial images
oNS	Only Natural language and Schematic images
oNMP	Only Natural language, Mathematical notation, and Pictorial images
oNMS	Only Natural language, Mathematical notation, and Schematic images
oNPS	Only Natural language, Pictorial images, and Schematic images
NMPS	Natural language, Mathematical notation, Pictorial images, and Schematic images

Only the *existence* of the four semiotic resources is noted, and no distinction is made between tasks with few or many instances of a particular semiotic resource. The difference in the number of instances of a particular semiotic resource between tasks is, of course, also an important factor worth focusing on, but in the current study it has been necessary to disregard such differences in order to enable a thorough analysis of all possible combinations of semiotic resources. To strengthen the reliability of the analysis of the tasks, every ambiguous categorisation is evaluated together with similar cases to find general rules that can guide the categorisation. Intra-rater reliability is achieved by coding the presence of the different semiotic resources in all tasks twice, with a six-month time span between the occasions. For all 497 analysed tasks, there were fewer than ten tasks for which the category was changed when the tasks were categorised a second time.

Two types of measures are used for how hard the tasks are to read and solve, namely *difficulty* and *demand on reading ability* (explained in next section). For both the SweNT and the PISA test, proportions of correct solutions are used to calculate a measure of task difficulty. A higher value stands for a more difficult task (i.e. difficulty = 1 – proportion of correct solutions).

Statistical analyses

In order to analyse if non subject-specific demand on reading ability (DRA) is related to some aspect of the multisemiotics of the mathematics tasks, values for the DRA are obtained through a *principal component analysis* (PCA) of all Swedish students' results on all PISA mathematical literacy tasks and reading literacy tasks. A PCA is a statistical method that, based on relations between existing variables in a dataset, extracts new underlying components from the same dataset. In a PCA, the components are constructed in such a way that the first principal component explains as much of the variation in the data as possible, and each subsequent component explains as much of the remaining variation as possible (see Tabachnick & Fidell, 2006). An oblique rotation is used in the PCA because the components are expected to correlate. It is reasonable to assume that several different aspects explain the students' results on the literacy and mathematics tasks and that a kind of reading ability would be the main factor explaining the results on the reading literacy tasks and that a mathematical ability would be the main factor explaining the results on the mathematics tasks. The two first components from the PCA are expected to correspond to the two abilities of mathematics and reading, an assumption strengthened by the pattern for how the analysed tasks are related to the components through the loading values. Each of the analysed tasks receives loading values for each of the components, and the loading values on the reading ability component are interpreted as a measure of the genuine effect of reading ability (the DRA) when the effect of mathematical ability has been excluded since the loading values represents the unique variance for a task explained by that component. This means that for mathematics tasks DRA can be seen as an *unnecessary* demand on reading ability (see also Dyrvold et al., 2015). Because access to the same students' results on both reading tasks and mathematics tasks are necessary for the analysis of the tasks' DRA values, the analysis has only been possible to conduct on the PISA tasks. For the measures of difficulty and DRA on the sample of PISA tasks, values are first obtained for the separate test years of 2003 and 2012. When the same task is used in both years, the mean value from the two years is used in the analyses.

Pearson correlation is used to test if the number of semiotic resources present in a task is related to difficulty or DRA, and a *p*-value less than .05 is considered significant. An independent samples *t*-test with .05 as the significance limit, is used to test whether there is any significant difference in difficulty or DRA between groups of tasks with and without a particular semiotic characteristic. For all *t*-tests the variance can be assumed to be equal for every pair of groups, according to Levene's test for equality of variances. Therefore the *t*- and *p*-values reported are the values obtained if equal variance is assumed. The *t*-test is two-tailed because no assumption is made regarding which of the groups tested might be more demanding to read or to solve. In some of the *t*-tests, there is a big difference between the size of the groups, and the smallest group might be very small. However, statistical analyses revealed that a *t*-test is reliable for analyses with extremely small samples (that is $n \geq 5$, divided into two groups) also when the group sizes are unequal (de Winter, 2013). Therefore, results where the particular semiotic characteristic is present in as few as five tasks are also valuable to take into account in the interpretation of the analysis.

Results

At first some descriptive statistics that enlighten some of the differences between the data samples are presented. In the next two sections the results from the statistical tests conducted to be able to answer the research question is presented. In tables 4–6 results that regards if the presence or co-occurrences of different semiotic resources are related to task difficulty or demand on reading ability is presented. In table 7 results that regards if the number of semiotic resources are related to task difficulty or demand on reading ability is presented. For a more explicit answer to all parts of the research question, see Conclusions.

Descriptive statistics

There are some differences regarding the number and percentage of tasks with a particular semiotic characteristic (the variables analysed) between PISA mathematics and the SweNT in mathematics (table 2). For three semiotic characteristics there are a large difference between the samples, namely tasks with schematic images (S), tasks with only natural language, mathematical notation, and schematic images present (oNMS),

Table 2. *Percentage of tasks with each semiotic characteristic in the two test samples*

Characteristic	PISA math * (%)	SweNT ** (%)
N	100	99.7
M	93.2	94.8
S	68.4	31.6
P	33.1	26.1
oN	0.8	1.1
oNM	14.3	49.5
oNS	3.8	2.5
oNP	0.8	0.8
oNMP	15.8	16.8
oNMS	48.1	20.6
oNSP	1.5	0.8
NMPS	15.0	7.7

Note. * The sample of tasks from PISA mathematics 2003 and 2012 (n = 133). ** The sample of tasks from the SweNT in mathematics from years 2004–2012 (n = 364).

which are more common in PISA, and tasks with only natural language and mathematical notation (oNM), which occur to a higher frequency in SweNT. All variable names are explained in table 1.

In the analysis of the PISA tasks, different samples are used in the analysis with difficulty (sample A) and with DRA (sample B, a subsample of A) as the dependent variable. The reason for this is that 28 tasks are excluded from the analysis with DRA (explained in Method), and thus sample B is not a random sample from A. This difference between the samples is also evident through the percentage of different multisemiotic task types in the samples. The percentage of tasks in the different samples is presented in table 3.

Relations between aspects of difficulty and semiotic characteristics

T-tests are used to examine whether there are any differences regarding the mean difficulty or DRA between groups of tasks with a particular semiotic characteristic and tasks without that characteristic. Only results for which there are at least five tasks in the smallest group are presented.

For the SweNT in mathematics, there are no significant differences in mean difficulty between groups of tasks with or without a particular semiotic characteristic. Results from the statistical tests conducted on the SweNT sample are presented in table 4.

For the tasks used in PISA 2003 and 2012, there are four semiotic characteristics for which the group of tasks significantly differ in mean difficulty depending on whether they have that semiotic characteristic or

Table 3. *Percentage of tasks with each semiotic characteristic in sample A and sample B from PISA 2003 and 2012 and in the group of excluded tasks*

Characteristic	Sample A: PISA math 2003 & 2012 (*n* = 133)	Sample B: PISA math 2003 & 2012 after exclusion of tasks (*n* = 105)	PISA math,the group of excluded tasks (*n* = 28)
N	100	100	100
M	93.2	92.4	96.4
S	68.4	71.4	57.1
P	33.1	26.7	57.1
oN	0.8	1.0	0.0
oNM	14.3	15.2	10.7
oNS	3.8	3.8	3.6
oNP	0.8	1.0	0.0
oNMP	15.8	11.4	32.1
oNMS	48.1	53.3	28.6
oNSP	1.5	1.9	0.0
NMPS	15.0	12.4	25.0

Table 4. *Difference in mean difficulty between tasks with or without a particular semiotic characteristic (SweNT)*

Grouping variable tested for difficulty	*n*	M	SD	*t*(364)	*p*-value	eta^2
M	345	0.464	0.213	0.132	0.895	0.00
No M	19	0.457	0.249			
S	115	0.477	0.221	0.853	0.394	0.00
No S	249	0.457	0.212			
P	95	0.485	0.220	1.167	0.244	0.00
No P	269	0.455	0.212			
oNM	180	0.453	0.211	-0.905	0.366	0.00
Not oNM	184	0.473	0.218			
oNS	9	0.450	0.279	-0.186	0.852	0.00
Not oNS	355	0.464	0.213			
oNMP	61	0.465	0.214	0.056	0.955	0.00
Not oNMP	303	0.463	0.215			
oNMS	75	0.468	0.210	0.218	0.828	0.00
Not oNMS	289	0.462	0.216			
NMPS	28	0.528	0.230	1.664	0.097	0.01
Not NMPS	336	0.458	0.213			

Table 5. *Difference in mean difficulty between tasks with or without a particular semiotic characteristic (PISA 2003 and 2012)*

Grouping variable tested for difficulty	*n*	M	SD	*t*(364)	*p*-value	eta^2
M No M	124 9	0.52 0.52	0.23 0.24	−0.025	0.98	0.000
S No S	91 42	0.50 0.55	0.24 0.21	1.143	0.255	0.010
P No P	44 89	0.59 0.48	0.23 0.22	−2.744	**0.007**	0.054
oNM Not oNM	19 114	0.48 0.52	0.20 0.24	0.699	0.486	0.004
oNS Not oNS	5 128	0.63 0.51	0.23 0.23	−1.12	0.265	0.009
oNMP Not oNMP	21 112	0.62 0.50	0.22 0.23	−2.132	**0.035**	0.034
oNMS Not oNMS	64 69	0.47 0.57	0.23 0.22	2.501	**0.014**	0.046
NMPS Not NMPS	20 113	0.61 0.50	0.23 0.23	−2.009	**0.047**	0.030

Table 6. *Difference in mean demand on reading ability (DRA) between tasks with or without a particular semiotic characteristic (PISA 2003 and 2012)*

Grouping variable tested for DRA	*n*	M	SD	*t*(364)	*p*-value	eta^2
M No M	97 8	0.245 0.222	0.124 0.086	0.522	0.603	0.003
S No S	75 30	0.236 0.261	0.120 0.125	−0.947	0.346	0.009
P No P	28 77	0.240 0.244	0.128 0.120	−0.175	0.861	0.000
oNM Not oNM	16 89	0.260 0.240	0.126 0.121	0.616	0.539	0.004
oNMP Not oNMP	12 93	0.263 0.240	0.141 0.119	0.608	0.545	0.004
oNMS Not oNMS	56 49	0.244 0.242	0.121 0.122	0.122	0.903	0.000
NMPS Not NMPS	13 92	0.211 0.248	0.124 0.121	−1.027	0.307	0.010

not (table 5). First, tasks that have pictorial images (P) are more difficult than tasks that have no pictorial images (No P). Second, tasks that have natural language, mathematical notation, and pictorial images, but no schematic images (oNMP), are more difficult than the rest of the tasks (Not oNMP). Third, tasks with all four semiotic resources (NMPS) are more difficult than tasks with at the most three semiotic resources (Not NMPS). Fourth, tasks that have natural language, mathematical notation, and schematic images, but no pictorial images (oNMS), have a lower mean difficulty than tasks that do not belong to that group. All four groups differ significantly in mean difficulty from tasks without those particular semiotic characteristics, and the effect size is small to moderate (between .030 and .054) (see e.g. Cohen, 1988, regarding effect size).

For the subsample (B) of tasks from PISA 2003 and 2012, there are no significant mean differences in DRA for groups of tasks with or without a particular semiotic characteristic. Results from the statistical tests are presented in table 6.

Relations between difficulty and different semiotic resources

For each of the three samples (one for SweNT and two for PISA) used in the current study, the relation between the number of different semiotic resources present in the tasks and the task's difficulty and DRA, is investigated using Pearson correlation. The statistical tests resulted in no significant correlations (table 7) and therefore the results do not give evidence for a relation between number of different semiotic resources in the tasks and how difficult the tasks are to read and solve.

Conclusions

The following section presents the conclusions drawn from the results regarding all three parts of the research question as well as the conclusions drawn based on all parts of the research question interpreted together.

Table 7. *Correlations between the number of different semiotic resources present and the difficulty or the demand on reading ability (DRA) for three samples of tasks*

Test sample	Measure	Measure: number of different semiotic resources present in tasks		
		Correlation coefficient	*p*-value	*n*
SweNT	Difficulty	.081	.125	364
PISA math (A)	Difficulty	.106	.225	133
PISA math (B)	DRA	−.061	.539	105

The answer to the research question – whether there is any relation between a task's semiotic characteristic and the difficulty or DRA of the task – is that there is a relation between four particular semiotic characteristics and how difficult the tasks are to solve, but there is no relation between any particular semiotic characteristic and DRA. Important to note, however, is that no group of tasks with a particular semiotic characteristic has a mean difficulty that differs significantly from the mean for the rest of the tasks in *both* PISA and SweNT. Pictorial images stand out because for all significant results pictorial images are present in the group of more difficult tasks, or are not present in the group of less difficult tasks. Also, there is no relation between the number of different semiotic resources in tasks and the difficulty or DRA of the tasks. These results, of course, only inform us about the presence of particular semiotic resources in tasks with a higher mean difficulty than the rest of the sample, and nothing can be said about causality.

When the results from the statistical tests are interpreted together, three additional conclusions can be drawn. First, pictorial images are a common factor in tasks that are less frequently solved correctly. For the PISA sample, there are three semiotic characteristics for which the tasks have a mean difficulty that is significantly higher than for the group of tasks without those features, namely tasks with pictorial images (P), tasks with only natural language, mathematical notation, and pictorial images (oNMP), and tasks with all four semiotic resources (NMPS). For tasks with only natural language, mathematical notation, and schematic images (oNMS), the tasks are significantly less difficult to solve. A common factor for these results is that pictorial images are present in the group of tasks that are more difficult. Because semiotic characteristic ”P” refers to all tasks with pictorial images irrespective of other semiotic resources being present, one possible explanation for the higher mean difficulty for that group is that the group includes two groups of tasks that are significantly more difficult (i.e. oNMP and NMPS).

Second, besides the lack of correlation between the number of different semiotic recourses in tasks and neither difficulty, nor DRA, the results give additional information. Based on the significant difference in mean difficulty between groups of tasks with different semiotic characteristics, it is not possible to draw any conclusions about some limit over which the number of different semiotic resources present in the task text are related to difficulty. The number of different semiotic resources in the task types that include pictorial images and have a significantly higher mean difficulty than the compared group varies from three to four different semiotic resources, but because there are too few tasks containing only pictorial images together with natural language (oNP) in the sample, that group is not tested for differences in mean difficulty in the PISA samples. Also, one group of tasks with three different semiotic resources (oNMS) have a significantly *lower* mean difficulty. Therefore, it is not necessarily the case that a mathematical task with many different semiotic resources will be more difficult to solve.

Third, there is no significant relationship between the task's DRA and its particular semiotic characteristic. This lack of a relationship between the task's DRA and any of the semiotic characteristics contributes to the understanding of what the differences in difficulty between tasks with or without a particular semiotic characteristic mean. Because groups of tasks do not differ in mean DRA depending on whether they fulfil the criteria for one of the four semiotic characteristics for which the group of tasks have a significantly different mean difficulty (P, oNMP, oNMS and NMPS), the difficulty aspect related to those semiotic characteristics is likely to be a mathematics-specific difficulty. This is because difficulty related to a non-mathematics-specific reading ability is expected to result in significant differences in mean DRA between groups of tasks with or without a particular semiotic characteristic (see also Dyrvold et al., 2015).

Discussion

In the interpretation of the results, some aspects of the method are important to take into account. Concerning the statistics, the number of *t*-tests performed necessarily means that the results must be interpreted with caution, particularly if conclusions are drawn based on single differences. This is because many statistical tests increase the risk that the analysis will result in significant differences stemming from coincidence. Also, with the decision to focus on the presence and co-occurrence of semiotic resources it follows that variation in the amount of the various semiotic resources in the tasks is disregarded in the analysis (i.e. the extent to which every semiotic resource is present). The decision to perform the analysis in this way was made despite this limitation because it allowed for all possible combinations of semiotic resources in the tasks to be analysed. Still, when the results are interpreted it is important to reflect on the variations within groups of tasks with particular semiotic characteristics.

The difference between the results of the *t*-tests for the groups of PISA tasks and the SweNT tasks is a bit unexpected because if tasks with a particular semiotic characteristic are strongly related to task difficulty, the *t*-test would result in a significant difference in mean difficulty despite minor differences between the two tests. However, there are several aspects of the SweNT that indicate the presence of more task features related to difficulty in the SweNT than in PISA and that these vary between tasks. For example, the tasks in the SweNT span from tasks demanding one calculation, giving a number as the answer, to a task that the students are expected to spend 50 minutes solving. Also, the SweNT uses images relating to a theme for an entire section of the test. Such features can make the tasks in a sample more heterogeneous and can also make a relation between tasks with a particular semiotic characteristic and difficulty more difficult to detect.

Earlier research has found that schematic images are difficult to decode (Lowrie et al., 2011) and that tasks with schematic images are more demanding than similar tasks without such images (Elia et al., 2007; Gagatsis & Elia, 2004; Lin et al., 2013). Based on these studies, it is reasonable to expect semiotic characteristics where schematic images are present to be more difficult, which in the current study is true only for the group of tasks with all four semiotic resources present. The relation between the presence of several semiotic resources in the task text and the difficulty of the task can be concluded from several studies. For example, it has been shown that it is difficult to translate between and to organize information from several semiotic resources (Ainsworth et al., 2002; Duru & Koklu, 2011). In the current study, the presence of multiple semiotic resources is a factor related to difficulty, but the crucial factor is not the number of different semiotic resources present, but *which* semiotic resources are present. This result suggests that the difficulty aspect related to the presence of multiple different semiotic resources in mathematical tasks is related to the traits of the semiotic resources that are present and how they interact. However, further research is needed regarding what the co-occurrence of different semiotic resources means for the reading and solving of the tasks.

It is a bit unexpected that instead of schematic images, it is pictorial images that recur in categories of tasks with a higher mean difficulty than others tasks. This indication that pictorial images are part of some aspect of difficulty might have to do with redundant information (as in the study by Berends & van Lieshout, 2009), particularly because some of the pictorial images have illustrative roles in the test. Altogether, the results of the current study contribute to the understanding of difficulty in relation to the semiotic characteristics of tasks by demonstrating an aspect of difficulty related to pictorial images and by showing that some aspects of difficulty are related to interactions between particular semiotic resources. Also, the results regarding DRA is important, since the nonexistence of significant relations to DRA indicate that the difficulty related to particular combinations of semiotic resources is a mathematics specific difficulty. Thus the type of difficulty that mathematics tests are designed to assess. Those results have implications for task design since a test with high validity assesses mathematical proficiency, nothing else.

The current study has pointed to the role of pictorial images on task difficulty, results that are of importance for the field of mathematics education research as one step against an understanding of how the semiotic feature of tasks might influence the reading and solving of a task. Reading pictorial images must not be perceived as trivial. The results of the current study do however not make it clear what the presence of other semiotic resources together with pictorial images means for the difficulty of a mathematical task. Studies focusing on the role of pictorial images in mathematical tasks can further our understanding of what the difficulty aspect related to pictorial images means.

References

Adams, T. L. (2003). Reading mathematics: more than words can say. *Reading Teacher*, 56 (8), 786–795.

Adams, T. L. & Lowery, R. M. (2007). An analysis of children's strategies for reading mathematics. *Reading & Writing Quarterly*, 23 (2), 161–177.

Ainsworth, S., Bibby, P. & Wood, D. (2002). Examining the effects of different multiple representational systems in learning primary mathematics. *Journal of the Learning Sciences*, 11 (1), 25–61.

Berends, I. E. & van Lieshout, E. M. (2009). The effect of illustrations in arithmetic problem-solving: effects of increased cognitive load. *Learning and Instruction*, 19 (4), 345–353.

Bergqvist, E. & Österholm, M. (2010). A theoretical model of the connection between the process of reading and the process of solving mathematical tasks. In C. Bergsten, E. Jablonka & T. Wedege (Eds.), *Mathematics and mathematics education: cultural and social dimensions. Proceedings of MADIF 7: the seventh Mathematics Education Research Seminar* (pp. 47–57). Linköping: SMDF.

Cohen, J. W. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale: Lawrence Erlbaum Associates.

Cromley, J. G., Snyder-Hogan, L. E. & Luciw-Dubas, U. A. (2010). Cognitive activities in complex science text and diagrams. *Contemporary Educational Psychology*, 35 (1), 59–74.

Duru, A. & Koklu, O. (2011). Middle school students' reading comprehension of mathematical texts and algebraic equations. *International Journal of Mathematical Education in Science & Technology*, 42 (4), 447–468.

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

Dyrvold, A., Bergqvist, E. & Österholm, M. (2015). Uncommon vocabulary in mathematical tasks in relation to demand of reading ability and solution frequency. *Nordic Studies in Mathematics Education*, 20 (1), 101–128.

Elia, I., Gagatsis, A. & Demetriou, A. (2007). The effects of different modes of representation on the solution of one-step additive problems. *Learning and Instruction*, 17 (6), 658–672.

Gagatsis, A. & Elia, I. (2004). The effects of different modes of representations on mathematical problem solving. In M. J. Høines & A. B. Fuglestad (Eds.), *Proceedings of the 28th conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 447–454). Bergen: PME.

Garderen, D. van (2006). Spatial visualization, visual imagery, and mathematical problem solving of students with varying abilities. *Journal of Learning Disabilities*, 39 (6), 496–506.

Goldin, G. A. (1998). The PME working group on representations. *Journal of Mathematical Behavior*, 17 (2), 283–301.

Hegarty, M. & Kozhevnikov, M. (1999). Types of visual-spatial representations and mathematical problem solving. *Journal of Educational Psychology*, 91 (4), 684–689.

Hiebert, J. & Carpenter, T. (1992). Learning and teaching with understanding. In D. A. Grouws (Ed.), *Handbook of research on mathemaitics teaching and learning* (pp. 65–97). New York: Macmillan.

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

Lemke, J. L. (1998). Multiplying meaning: visual and verbal semiotics in scientific text. In J. R. Martin & R. Veel (Eds.), *Reading science* (pp. 87–113). London: Routledge.

Lemke, J. L. (2003). Mathematics in the middle: measure, picture, gesture, sign, and word. In M. Andersson, A. Sáenz-Ludlow, S. Zellweger & V. Cifarelli (Eds.), *Educational perspectives on mathematics as semiosis: from thinking to interpreting to knowing* (pp. 215–234). Ottawa: Legas Publishing.

Lin, Y., Wilson, M. & Cheng, C. (2013). An investigation of the nature of the influences of item stem and option representation on student responses to a mathematics test. *European Journal of Psychology of Education*, 28 (4), 1141–1161.

Lowrie, T., Diezmann, C. M. & Logan, T. (2011). Understanding graphicacy: students' making sense of graphics in mathematics assessment tasks. *International Journal for Mathematics Teaching and Learning*, 12, 1–32. Retreived from http://www.cimt.plymouth.ac.uk/journal/default.htm

Martiniello, M. (2009). Linguistic complexity, schematic representations, and differential item functioning for English language learners in math tests. *Educational Assessment*, 14 (3-4), 160–179.

Mitchell, W. (1986). *Iconology: image, text, ideology*. University of Chicago Press.

Mundy, E. & Gilmore, C. K. (2009). Children's mapping between symbolic and nonsymbolic representations of number. *Journal of Experimental Child Psychology*, 103 (4), 490–502.
NCTM. (2000). *Principles and standards for school mathematics*. Reston: National Council of Teachers of Mathematics.
Niss, M. (2003). *Mathematical competencies and the larning of mathematics: The Danish KOM project*. Paper presented at the third Mediterranean conference on mathematics education, Athens.
O'Halloran, K. (2005). *Mathematical discource: language, symbolism and visual images*. London: Continuum.
O'Halloran, K. (2007). Systemic functional multimodal discourse analysis (SF–MDA) approach to mathematics, grammar and literacy. In R. Wittaker, M. O'Donnel & A. McCabe (Eds.), *Advances in language and education* (pp. 77–102). London: Continuum.
O'Halloran, K. (2008). Inter-semiotic expansion of experiental meaning: hierarchical scales and metaphor in mathematics discourse. In E. Ventola & C. Jones (Eds.), *From language to multimodality. New developments in the study of ideational meaning* (pp. 231–254). London: Equinox.
Pimm, D. (1987). *Speaking mathematically: communication in mathematics classrooms*. London: Routledge Kegan & Paul.
Pimm, D. (1995). *Symbols and meanings in school mathematics*. London: Routledge.
Radford, L. & Puig, L. (2007). Syntax and meaning as sensous, visual, historical forms of algebraic thinking. *Educational Studies in Mathematics*, 66, 145–164.
Rousselle, L. & Noel, M. (2007). Basic numerical skills in children with mathematics learning disabilities: a comparison of symbolic vs non-symbolic number magnitude processing. *Cognition*, 102 (3), 361–395.
Sfard, A. (2008). *Thinking as communicating: human development, the growth of discources, and mathematizing*. Cambridge University Press.
Tabachnick, B. G. & Fidell, L. S. (2006). *Using multivariate statistics* (Vol. 5, rev. ed.). Boston: Allyn and Bacon.
Vukovic, R. K. & Lesaux, N. K. (2013). The language of mathematics: Investigating the ways language counts for children's mathematical development. *Journal of Experimental Child Psychology*, 115 (2), 227–244.
Winter, J. C. F. de (2013). Using the student's *t*-test with extremely small sample sizes. *Practical Assessment, Research & Evaluation*, 18 (10), 1–12.
Österholm, M. (2006). Characterizing reading comprehension of mathematical texts. *Educational Studies in Mathematics*, 63, 325–346.
Österholm, M. & Bergqvist, E. (2012). Methodological issues when studying the relationship between reading and solving mathematical tasks. *Nordic Studies in Mathematics Education*, 17 (1), 5–30.

Anneli Dyrvold

Anneli Dyrvold is a PhD student in mathematics education at the Department of Mathematics and Mathematical Statistics at Umeå University. She is a member of Umeå Mathematics Education Research Centre (UMERC). Her research interest is the affordances, and distinguishing characteristics, of the language of mathematics and what students need to learn in order to master it.

anneli.dyrvold@umu.se