Eur J Psychol Educ (2018) 33:505–520
https://doi.org/10.1007/s10212-018-0385-x

Italian students' results in the PISA mathematics test: does reading competence matter?

Anna Maria Ajello[1,2] · **Elisa Caponera**[1] · **Laura Palmerio**[1]

Received: 31 May 2017 / Revised: 13 April 2018 / Accepted: 19 April 2018 /
Published online: 15 May 2018
© Instituto Superior de Psicologia Aplicada, Lisboa, Portugal and Springer Science+Business Media B.V., part of Springer Nature 2018

Abstract In Italy, from the 2003 reports to the present, the National Institute for the Educational Evaluation of Instruction and Training (INVALSI) has conducted research on Programme for International Student Assessment (PISA) results in order to understand Italian students' low achievement in mathematics. In the present paper, data from a representative sample of 15-year-old Italian students who participated in PISA 2012 were analysed. This study's primary aim is to verify how students' linguistic competences are associated with their performance in mathematics. For the evaluation of the impact of item reading demand on students' performance, we selected 24 mathematics items with a high reading demand and 31 mathematics items with a low reading demand, as classified by Italian language and methodology experts. Repeated measure variance analyses were conducted. The results showed differences in function of gender: females are advantaged in mathematics items with a high reading demand, independent of their level of reading literacy. In contrast, males are advantaged in mathematics items with a low reading demand, independent of their level of reading literacy. Possible policy implications are discussed.

Keywords OECD PISA · Mathematics achievement · Reading · Gender differences

✉ Laura Palmerio
laura.palmerio@invalsi.it; http://www.invalsi.it

Anna Maria Ajello
presidente@invalsi.it; annamaria.ajello@uniroma1.it; http://www.invalsi.it; http://www.psicologia1.uniroma1.it/

Elisa Caponera
elisa.caponera@invalsi.it; http://www.invalsi.it

[1] INVALSI (National Institute for the Educational Evaluation of Instruction and Training), Via Ippolito Nievo 35, 00153 Rome, Italy

[2] Department of Psychology, Sapienza Università di Roma, Via dei Marsi, 78, 00185 Rome, Italy

Introduction

Over the past decades, mathematics has become increasingly important, and recent research (e.g. Nasser and Birebaum 2005) highlights that high mathematics achievers have more opportunities for personal growth and are more likely to find a job, build a better career and become active citizens. Consequently, identifying the strategies to improve mathematics achievement is a crucial issue for policymakers (European Commission 2010, 2012). In fact, one of the aims of the Programme for International Student Assessment (PISA), promoted by the Organisation for Economic Co-operation and Development (OECD), is to 'measure how effectively countries are preparing students to use mathematics in every aspect of their personal, civic and professional lives, as part of their constructive, engaged and reflective citizenship'. Mathematical competence is defined as 'an individual's capacity to formulate, employ and interpret mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts and tools to describe, explain, and predict phenomena. It assists individuals to recognise the role that mathematics plays in the world and to make the well-founded judgments and decisions needed by constructive, engaged and reflective citizens' (OECD 2013, p. 17).

Several studies investigated factors that can contribute to explaining students' achievement (e.g. Schmidt et al. 2015; McConney and Perry 2010; Marks et al. 2006), and in many countries, several initiatives have been promoted in order to reduce differences in the achievement results (e.g. Breakspear 2012; OECD 2010; Grek 2009). The experience from the large comparative international surveys, such as PISA, enabled participating countries to obtain an accurate and reliable picture of the efficacy of their school systems.

Italian students achieved disappointing results in mathematics in almost all of the PISA cycles and had substantial difficulties using mathematical knowledge to solve the problems of everyday life. Table 1 shows the Italian results in PISA mathematics achievement since the 2000 cycle compared to the OECD mean. Even if the trends show a growth of our mathematics scores, the results remain below the OECD mean across the cycles, except for the 2012 cycle, in which Italian students have reached the OECD mean.

Table 1 Italian trends in PISA mathematics results

Cycle year	2003		2006		2009		2012	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Italy	466	96	462	96	483	93	485	93
OECD mean	500	100	498	92	499	92	494	92

In Italy, the National Institute for the Educational Evaluation of Instruction and Training (INVALSI) has been the National Centre for OECD PISA and for the International Association for the Evaluation of Educational Assessment (IEA) surveys since 1985. Since the PISA 2003 report to date, INVALSI has been conducting research aiming to understand Italian student achievement in mathematics (e.g. Arzarello et al. 2015). The research aimed to answer such general questions like 'What are the factors that influence the difficulty of PISA mathematics items?' or 'How does the complexity of the language used in the items serve as an indicator of their difficulty and more specifically, is it possible that failure in mathematics is partly due to the interference of the text elements used to convey the mathematical problems to be solved?'

To start answering the questions described above, investigations have been conducted by the INVALSI teams during the last 15 years. For example, a first study (Pozio 2011) was conducted using Italian data from PISA 2003 mathematics tests, divided into two steps:

1) One thousand booklets belonging to students, representative of two types of school (technical and vocational - artistic), from all of the Italian macro-areas were selected. A database of the observed errors was compiled, and errors related only to the released items were then examined.
2) The second phase focused on an administration of the tests to 40 15-year-old students in Rome who were attending technical or vocational schools. The students were interviewed as they were taking the test (i.e. 'think aloud' protocols were used) as to evaluate their reactions to this type of questions and to observe strategies and cognitive processes they used while solving the items.

The results showed that a limited knowledge of particular topics presented in the tests at times was often coupled with a low comprehension of the items' written text or a misinterpretation by the students, probably caused by a lack of attention or a poor reading ability.

Thus, reading ability seemed to play a crucial role in mathematics learning and results. Reading, in fact, is a specific ability a student must acquire at school and, at the same time, it is the natural vehicle for any other school subjects. Reading literacy has been defined as a cross-curricular competency that consequently will affect achievement in other school subjects (e.g. Artelt et al. 2001). Mathematics tasks that are usually used to assess students' learning at school commonly consist of a text written in natural language plus mathematical notation and different types of images. Hence, being able to read and interpret such texts is an essential part of mathematical proficiency as it is transferred in the school context.

Although there is a consensus on the matter, the role concretely played by reading comprehension in learning other disciplines is still under debate (e.g. Anstrom et al. 2010; Baker 1991). For example, several studies (Neufeld 2006; Duke and Pearson 2002) showed that reading literacy implies a bias, as the higher the proficiency in this competency is, the greater is the advantage in learning other subjects. 'Students with high reading competence' have a richer vocabulary and therefore can better understand new texts, disregard irrelevant information and fill gaps in comprehension. Furthermore, previous research observed a positive relationship between standardised tests in mathematics and reading skills (Caponera et al. 2016; Mullis et al. 2013; Jiban and Deno 2007; Thurber et al. 2002; Helwig et al. 1999; Aiken 1972). Grimm (2008), for example, observed that the level of reading comprehension at the third grade was a positive and significant predictor of mathematics skills in eighth-grade students, after controlling for prior mathematics skills and student characteristics,

especially in regard to tasks involving the application of mathematics knowledge. The relation was stronger with the Mathematical Concepts and Estimation test than with Mathematical Computation. In a sample observing primary students from grade one to grade three, Hickendorff (2013) also reported that reading was strongly associated with mathematics results and that this relationship was stronger with applied problem-solving than with computation skills.

The use of language and symbols in mathematics becomes more demanding with each increasing grade level. By secondary school, reading skills are an even more powerful predictor of math performance (Korhonen et al. 2012). For example, Österholm (2006), comparing tasks with the same mathematical content, but with or without mathematical notation, observed that the presence of mathematical symbols in the text made it more difficult. The author underlines how reading a mathematical text is in itself not seen as an opportunity for learning but only as a necessary kind of ability in order to become active in situations where learning can occur (i.e. solving given tasks). Conversely, Driver and Powell (2015) investigated the difference between second-grade students' scores on mathematics tasks with and without mathematical notation. These researchers used a sample of students with/or at risk of mathematics difficulties. Both students with and without mathematics difficulties scored significantly lower on the tasks with the mathematical notation (Driver and Powell 2015).

Furthermore, the results from the previous study by Abedi (2000) highlighted that when the text in the tasks is simplified, the tasks are more likely to be solved correctly by students less proficient in the language of the test. The author conducted a study using data derived from the mathematics assessment of the National Assessment of Educational Progress (NAEP). The results clearly indicated an effect of student language backgrounds on their mathematics results in junior high school. Taking into account the length of items, students who spoke a language other than English at home performed significantly worse than students who spoke English at home, with a greater difference being observed on long items. Furthermore, the rates of omitted and not-reached math items were higher for non-native English speakers than those for native English speakers. That finding means that math performance is confounded with language performance on NAEP items and that students' language background impacts their performance.

Mathematics teachers who integrate texts into mathematics recognise that mathematical understanding involves reading and writing. Conversely, mathematical reasoning and problem-solving can be found in authentic reading and writing materials (Moyer 2000). Nonetheless, when a student is not successful in mathematics, teachers frequently assume the difficulty lies in the student's mathematical ability or negative disposition towards mathematics, but the difficulty may lie with the student's reading comprehension (Draper et al. 2005; Kane et al. 1974).

Several studies, for example, have underlined the importance of incorporating language teaching into specific content area instruction (e.g. mathematics) to improve student content area learning (Yore and Treagust 2006; Draper 2002). For example, Yore et al. (2007) highlighted that it is important to switch from home language to school language onto mathematical language to improve the students' competencies in this area. The National Council of Teachers of Mathematics (NCTM), a mathematics education organisation throughout the USA and Canada, noted that different verbal languages, symbol systems, gestures and representations play a relevant role in the production of mathematics (NCTM 2000). Integration of content areas has been studied for a number of years (Moyer 2000) and is promoted by many educators because of its benefits for learners of all ages and backgrounds (Burns 2005; Douville et al. 2003; Pica 2002). Draper (2002), referred to mathematics and literacy as 'inseparable' in constructivist classroom settings. Incorporating content area reading strategies into classroom instruction may be a vehicle through which teachers can facilitate students' ability to learn from their mathematics textbooks (Siebert and Draper 2008; Snow 2002).

Furthermore, if one considers that mathematics has its own specific vocabulary and that teaching mathematical topic requires verbal language, it is clear that improving students' linguistic skills is crucial in ensuring mathematics learning. It should be noted that a student who is faced with an unclear language could fall back on a mnemonic repetition of the topic with no real understanding (for details, see Borsese et al. 2014).

The evidence derived from international studies is confirmed in the Italian context: recent research highlighted that since kindergarten, certain difficulties in solving mathematical problems are due to poor linguistic skills (e.g. D'Amore 2014). Boero et al. (2008) underlined that natural language is crucial in learning mathematics, particularly in relation to the process of argumentation. Bolondi and Viale (2014) noted that on the one hand, linguistic skills are recognised as an essential factor by mathematics teachers, but on the other hand, this aspect is disregarded by the same teachers when assessing student achievement.

In addition, the results of the international surveys, such as PIRLS and PISA, show that girls outperform boys in reading literacy in most countries (OECD 2010; Stoet and Geary 2013) and from the primary school through upper secondary school. A recent study on PISA data in the USA shows that gender differences that favour boys are more noticeable for complex multiple items (Liu and Wilson 2009). Similarly, the IEA's Trends in International Mathematics and Science Study (TIMSS) 2011 shows that grade four boys performed better than the girls in math in most European countries, while no significant differences were found for the eighth grade. Previous cycles in both PISA and TIMSS have consistently indicated boys to be outperforming girls in mathematics across a number of countries (Education, A Culture Executive Agency 2010).

Several international and national studies documented systematic differences in the mathematics results of Italian boys and girls in favour of the former (e.g. INVALSI 2016; Mullis et al. 2016; OECD 2016). In Italy, differences between boys and girls are always in favour of boys and this trend is more consistent than in the other participating countries (OECD 2016; Mullis et al. 2016). It should also be noted that such differences tend to increase with the level of student education and thus the complexity of the texts.

In connection to gender differences, Marks (2008) argues that gender differences are related to school system organisation, students' expectations and environmental characteristics. Furthermore, for the author, some aspects of reading items (i.e. question format, content domains and cognitive domain) are closely associated with gender difference in student performance (for details, see Lafontaine and Monseur, 2009).

In mathematics, gender differences seem to be related to cognitive processes investigated by the question and linked to the type of question. While boys outperform girls in multiple-choice questions, girls outperform boys on open-ended questions (Bolger and Kellaghan 1990; Wilder and Powell 1989).

The previously discussed literature leads to the hypothesis that the reading burden of mathematics items is associated with student performance in mathematics, independent of their mathematics ability. Based on previous evidence, this study focused on the role played by students' reading literacy when dealing with the mathematics items. Consequently, the aim of the study was to explore the following:

1. What is the relationship between mathematics achievement and reading achievement in PISA 2012 in Italy;
2. Whether and to what extent the reading difficulty of mathematics items is associated with students' mathematics achievement; and
3. Whether and to what extent the differences in mathematics achievement between students with high reading competence and students with low reading competence is related to the student's gender.

Method

Participants

In the present study, analyses were conducted on the Italian PISA 2012 data. In Italy, 31,073 students, grouped in 1194 schools, participated in PISA 2012. Cases with missing values in one or more explanatory variables were excluded from the analyses. The overall sample consisted of 30,154 (49% girls), clustered in 1194 schools, representative of Italian 15-year-old students. All schools that participated in PISA 2012 were included in the analysis and the overall exclusion rate was 3%. The exclusion rate did not affect the sample composition.

Measures

The scales used in this study are described below.

Reading achievement scale The scale consisted of paper-based multiple-choice and constructed-response items. Reading achievement scale measured the capacity to understand, use and reflect on written texts. The whole item pool was 44; various combinations of the assessment items were assembled into 9 booklets.

Mathematics achievement scale Mathematical achievement scale measured the capacity to formulate, employ and interpret mathematics in a variety of contexts to describe, predict and explain phenomena and recognising the role that mathematics plays in the world. The scale consisted of 110 paper-based multiple-choice and constructed-response items. A cluster rotation design was used to form 13 booklets.[1] Since mathematics literacy was the major subject area of PISA 2012, the International PISA Consortium computed a mean score for the combined mathematics literacy scale, as well as a mean score for each of the three domains: content (personal, occupational, societal and scientific), context (change and relationships, space and shape, quantity, and uncertainty and data) and processes (formulating, employing, interpreting).

Each student responded to only one booklet containing both mathematics and reading items. Using IRT estimates, a score of reading achievement and a score of mathematics achievement were calculated for each student. To take into account measurement errors, a range of five plausible values of scores for each domain was provided for each student (for a detailed description, see OECD 2014a).

Data analysis

In this study, analyses were performed using the software SPSS 20, by means of adapted macros provided by the PISA Consortium. Correlation analyses for Italian students were conducted to verify the relation between the two domains.

For the evaluation of the impact of item reading demand on students' performance, we selected 24 mathematics items with a high reading demand and 31 mathematics items with a low reading demand. The classification was conducted by Italian methodology and subject experts based on recommendations derived from international studies (e.g. Mullis et al. 2013).

To classify the items, we considered the natural language (words and letters), the types of images (graphs and tables) and the mathematical notation (the mathematics symbolic language and specialised vocabulary) used in the items, and we referred to the criteria described in Table 2.

[1] In PISA 2012, each of 13 booklets included a subset of mathematics items, while only 9 out of these 13 booklets included reading items.

Table 2 Criteria for item classification

	Type of item	
	Low demand	High demand
Number of words	Below 100 words	Above 200 words
Use of symbols	No	Yes
Specialised vocabulary (lexicon)	No	Yes
Presence of graphs and tables	No	Yes

An item meeting criterion 1 (with more than 200 words) and at least two out of the three other criteria with high demand was classified as a high reading demand item; an item with less than 100 words and low demand with regard to at least two out of the three other criteria was classified as a low-reading-demand item. The remaining items were classified as medium-demand items. Some sample items with low and high reading demands are presented in the Appendix.[2]

For each student, we computed two separate standardised scores averaging their performance in these two categories of items, one for high-reading-demand items and one for low-reading-demand items. Using Rasch estimates, we calculated two separate scores of mathematics achievement for each student. For each of these two categories, separately, we computed a set of five plausible values of scores for each student to take into account measurement errors depending on the fact that not all items were administered to all students. Next, we converted these scores into two standardised scores with a scaled average of 500 and standard deviation of 100.

After that step, we selected a subsample of students based on their level of reading ability on the PISA reading test. Students with low reading competence were defined as those at levels 1a and 1b on proficiency levels on the reading literacy scale, and students with high reading competence were defined as those at levels 5 and 6 on proficiency levels on the reading literacy scale (for a description of proficiency levels, see OECD 2014b).

Next, we performed a mixed analysis of variance, with item reading demand (two levels: high vs. low) as within factor and reading literacy (two levels: students with low reading competence/students with high reading competence) and gender as between factors.

Results

Descriptive statistics

Table 3 shows the students' sample description and the descriptive statistics for reading and mathematics achievement for all students and for students of each gender.

The results showed significant gender differences: boys performed better in mathematics than girls (497 vs. 477), while girls outperformed boys in reading (511 vs. 474). Moreover, Italian students performed worse than the OECD International average in reading (493 vs. 500) and mathematics (487 vs. 500).

[2] The classification of the items used in this study has been performed on the Italian version. Nonetheless, to facilitate the reader, the English version of the sample items is included in the Appendix.

Table 3 Students' sample description and descriptive statistics for mathematics (all students and by gender)

	Participants		Reading		Mathematics	
	Sample	Pop. estimate	Mean	S.D.	Mean	S.D.
All students	30,154	505,837	493	91	487	89
Girls	14,900	246,457	511	83	477	84
Boys	15,254	259,380	474	95	497	93

The results indicated a significant difference in mathematics between two groups, students with low reading competence and students with high reading competence (Table 4). On average, students with high reading competence outperformed students with the low reading competence of two standard deviations and a half. Furthermore, the correlation between reading literacy and mathematics was quite high with a coefficient of 0.84, and 45% of students who reached a certain proficiency level in reading reached the same proficiency level in mathematics, as well. Table 5 shows results on mathematics performance for boys and girls separately and considering items with low reading demand and high reading demand.

Table 4 Descriptive statistics in mathematics achievement among students with low reading competence and students with high reading competence across gender

	Mathematics achievement			
	Students with low reading competence		Students with high reading competence	
	Mean	S.D.	Mean	S.D.
All students	381	55	629	52
Girls	354	46	612	48
Boys	393	54	661	42

Table 5 Students' mean score as a function of gender and reading proficiency level

Gender	Students with low reading competence			Students with high reading competence		
	Items with low reading demand	Items with high reading demand	Difference between low reading demand and high reading demand	Items with low reading demand	Items with high reading demand	Difference between low reading demand and high reading demand
Girls	355	364	*− 9*	578	588	*− 10*
Boys	407	397	*10*	633	628	*5*

Statistically significant differences are in italic, $p < 0.01$

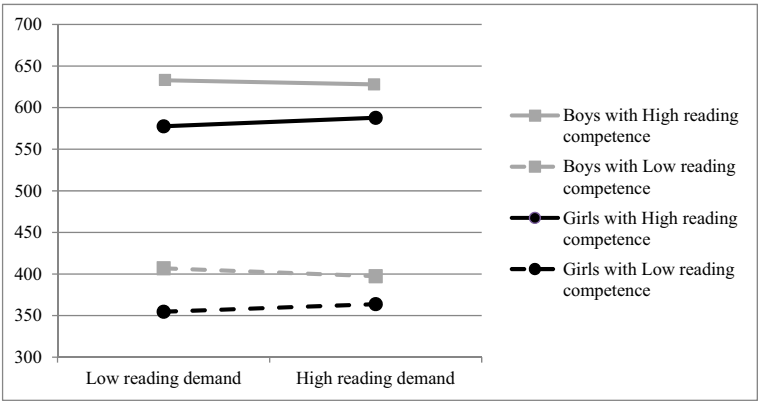


Fig. 1 Reading demand and mathematics achievement by gender

All differences are statistically significant. The mixed model ANOVA showed a significant difference both between students with high reading competence and students with low reading competence and between boys and girls, depending on the reading demand level of the mathematics test. Students with high reading competence performed better than students with low reading competence. Boys performed better on the low-reading-demand mathematics scale than the high-reading-demand mathematics scale, independent of their reading literacy level. Conversely, girls performed better in the high-reading-demand mathematics scale than the low-reading-demand mathematics scale, again independent of their reading literacy level.

The difference between students' scores in the same category (gender and level of reading) for items with low and high reading demand is statistically significant for all groups. More specifically, girls with low reading competence performed better on items with high reading demand ($F_{(1,27346)} = 1137$; $p < .001$) ($\eta^2 = .04$), same as the girls with the high reading competence ($F_{(1,17546)} = 1762$; $p < .001$) ($\eta^2 = .09$). In contrast, boys with low reading competence performed better on items with low reading demand ($F_{(1,62669)} = 2547$; $p < .001$) ($\eta^2 = .04$), while the boys with high reading competence performed also better on items with low reading demand ($F_{(1,10572)} = 172$; $p < .001$) ($\eta^2 = .02$) (Fig. 1).

Discussion

The primary aim of the present study has been to investigate the relationship between mathematics achievement and reading achievement in Italian students participating in PISA 2012 with respect to the reading difficulty of mathematics items.

Our results have indicated that the association between reading and mathematics tests was very strong (Caponera et al. 2016; Mullis et al. 2013; Jiban and Deno 2007). These results were aligned with evidence from the previous literature confirming that reading literacy could produce a bias because students with high reading competence are advantaged in learning any other subject (regardless of their specific ability in those subjects) (Neufeld 2006; Duke and Pearson 2002). Thus, the strong correlation between reading literacy and mathematics might represent a problem of the construct validity of the PISA standardised test and should be confirmed in other national and international studies.

However, this psychometric flaw could become a merit from the point of view of the teaching practice. In fact, the present study evidenced the relevance of two different aspects to explaining mathematics results: student reading ability on the one hand and text characteristics on the other hand. The use of mathematics problems situated in a real-world context could help students to better memorise and/or understand mathematics concepts. Nevertheless, further studies should be conducted to investigate the mechanisms underlying the comprehension process. In this direction, we should value the predictive validity of the PISA test, considering that the literature showed a strong connection between linguistic competence and mathematics in school learning (Bolondi and Viale 2014; Korhonen et al. 2012; Boero et al. 2008; Schleppegrell 2007).

Moreover, the results of mixed method ANOVA showed that the reading burden of mathematics items was associated with the performance in the mathematics test in different ways, depending on the students' gender. Girls are advantaged in mathematics items with a high reading demand, independent of their level of reading literacy. In contrast, boys are advantaged in mathematics items with a low reading demand. A possible explanation of these results could be related to the item difficulty and to the type of the tasks requested. These results should be confirmed by further studies considering also what type of process is necessary to apply to answer correctly, for example, if the question requested computation instead of applied problem-solving.

This could help to clarify how linguistic complexity in mathematics items may be associated with student performance in mathematics tests as a function of these particular skills. This question appears to be an important one, as growing attention in schools is given to reading comprehension within other subjects, and different subject areas are increasingly integrated into teaching and learning (Martin and Mullis 2013; Mullis et al. 2009; Johnson and Zabrucky 2011).

From the present study, it appears that girls benefit from the linguistic complexity of mathematics items. In Italy, boys outperform girls both in international and national standardised mathematics tests at upper secondary school (INVALSI 2016; OECD 2016), showing also that girls perceive themselves as having lower self-efficacy and to be more anxious than the boys in mathematics. This self-regulation system primarily derives from previous experiences but, at the same time, could have an impact on future performances, as well as study and career expectations (Marsh et al. 2005). In Italy alone, gender disparity is very much visible when students are enrolled in the university. As an example, in the academic year 2016/2017, 38% of freshmen in natural science courses were girls, and 62% were boys. Even though the number of girls enrolling at the university as a whole is greater than the number of boys, only 25% of girls freshmen choose to study courses in natural science compared to 51% of boys (MIUR 2017). A recent study also found evidence that writing activities could be a key to increasing female student interest in mathematics (Reilly 2014). Thus, the comprehension of the different ways in which boys and girls learn mathematics could be crucial to reduce the gender gap.

Conclusion

The central aim of the study was to understand the role played by linguistic competencies in mathematics achievement. The overall picture of this paper found evidence of the need to deeply understand what kind of strategies our students use to solve mathematics problems. This could help identify possible teaching strategies to improve student comprehension of mathematics problems and consequently guarantee to our students the acquisition of an authentic competence to face mathematics problems that they may encounter in their everyday life.

On this subject, further research is necessary to investigate the influence of readability of constructed-response items on student achievement in mathematics standardised tests, thereby helping

us understand the importance of reading literacy on learning other content areas (e.g. if and in which way certain textual features are related to task difficulty of mathematics tasks).

Different limitations of this study should be noted. First, as the study involves one school grade only, further studies using data from different school grades are warranted in order to ascertain the impact of reading literacy on mathematics achievement of Italian students. Second, the results of this work should be confirmed and examined more in-depth using qualitative methods in order to better understand whether girls apply different cognitive strategies than the boys in addressing mathematical tasks.

Being aware of these limitations, in the present study, we examined the relationship between reading literacy and mathematics results on a representative student sample, assessed through a well-known and acknowledged international standardised test. It should be noted that, in recent years, data derived from PISA and other international surveys, such as IEA studies, have been used more and more by policymakers to improve teaching and learning in mathematics and the quality of education systems. Italian students, girls especially, showed relatively poor results in the mathematics tests. The study presented in this paper has attempted to investigate certain factors that could explain such results.

These point us to the conclusion that readability of mathematics tests is a critical aspect to be considered if one wants to appropriately assess student achievement in this subject. Moreover, we have recently witnessed a growing interest from teachers in integrating reading into mathematics lessons. There is, in fact, a consensus regarding the need to use a wide range of teaching methods to motivate all students, especially those with a low proficiency level in mathematics. This study seems to confirm the close relationship between reading and mathematics and offers some prompts to policymakers regarding the need for specific interventions in this direction. Nonetheless, there is a lack of empirical research investigating in-depth the relationship between mathematics and reading, and consequently studying how reading could practically help the mathematics learning. This study is a small contribution in that direction.

Examples of low-/high-reading-demand items

SAUCE

Question 2: SAUCE PM924Q02 – 0

You are making your own dressing for a salad.

Here is a recipe for 100 millilitres (mL) of dressing.

Salad oil:	60 mL
Vinegar:	30 mL
Soy sauce:	10 mL

How many millilitres (mL) of salad oil do you need to make 150 mL of this dressing?

Answer: …….. mL

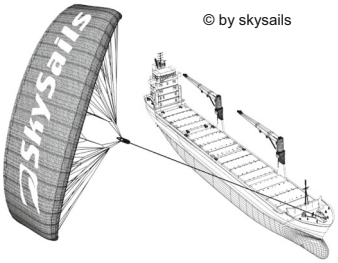
Fig. 2 Example of low-reading-demand item

SAILING SHIPS

Ninety-five percent of world trade is moved by sea, by roughly 50 000 tankers, bulk carriers and container ships. Most of these ships use diesel fuel.

Engineers are planning to develop wind power support for ships. Their proposal is to attach kite sails to ships and use the wind's power to help reduce diesel consumption and the fuel's impact on the environment.

Translation Note: "© by skysails": Do not adapt skysails as this is a registered label.



Question 1: SAILING SHIPS

PM923Q01

One advantage of using a kite sail is that it flies at a height of 150 m. There, the wind speed is approximately 25% higher than down on the deck of the ship.

At what approximate speed does the wind blow into a kite sail when a wind speed of 24 km/h is measured on the deck of the ship?

A 6 km/h
B 18 km/h
C 25 km/h
D 30 km/h
E 49 km/h

Translation Note: In this unit please retain metric units throughout.

SAILING SHIPS SCORING 1

QUESTION INTENT:

Description: Apply calculation of percentage within a given real world situation
Mathematical content area: Quantity
Context: Scientific
Process: Employ

Full Credit

Code 1: D. 30 km/h

No Credit

Code 0: Other responses.

Code 9: Missing.

Fig. 3 Example of high-reading-demand item

References

Abedi, J. (2000). Confounding of students' performance and their language background variables. Research Report 143, UD 033 934.

Aiken Jr., L. R. (1972). Language factors in learning mathematics. *Review of Educational Research, 42*(3), 359–385.

Anstrom, K., DiCerbo, P., Butler, F., Katz, A., Millet, J., & Rivera, C. (2010). *A review of the literature on academic English: implications for K-12 English language learners*. Arlington, VA: The George Washington University Center for Equity and Excellence in Education.

Artelt, C., Schiefele, U., & Schneider, W. (2001). Predictors of reading literacy. *European Journal of Psychology of Education, 16*(3), 363–383.

Arzarello, F., Garuti, R., & Ricci, R. (2015). The impact of PISA studies on the Italian national assessment system. In Stacey K. & Turner R. (Eds.), Assessing mathematical literacy. The PISA experience (pp.249–260). Springer International Publishing Switzerland.

Baker, L. (1991). Metacognition, reading and science education. *Science learning: processes and applications*, 2–13.

Boero, P., Douek, N., & Ferrari, P. L. (2008). Developing mastery of natural language: approach to theoretical aspects of mathematics. In L. D. English (Ed.), *Handbook of international research in mathematics education* (pp. 262–295). New York-London: Routledge.

Bolger, N., & Kellaghan, T. (1990). Method of measurement and gender differences in scholastic achievement. *Journal of Educational Measurement, 27*(2), 165–174.

Bolondi, G., & Viale, M. (2014). Abilità linguistiche e discipline scientifiche: un'esperienza di formazione del corpo insegnante nel Polo dell'Emilia-Romagna del progetto "I Lincei per una nuova didattica nella scuola" [Linguistic skills and scientific disciplines: an experience of teacher training in the Emilia-Romagna region in the project "The *Lincei* for a new didactics in the school"], XVIII Convegno nazionale GISCEL – Educazione linguistica e apprendimento/insegnamento delle discipline matematico-scientifiche. In *27–29 marzo*.

Borsese A., Mallarino B., Parrachino I., Rebella I. (2014). La definizione nel processo di insegnamento con particolare riferimento all'ambito scientifico [Definition in the teaching process with particular reference to the scientific field], XVIII Convegno Nazionale GISCEL – Educazione linguistica e apprendimento/insegnamento delle discipline matematico-scientifiche, 27–29 marzo. http://www.giscel.it/sites/default/files/convegni/2014/2014-03%20OPUSCOLO_PREPRINT%20Giscel.pdf Accessed 24/04/2017.

Breakspear, S. (2012). The policy impact of PISA: an exploration of the normative effects of international benchmarking in school system performance. *OECD Education Working Papers*, (71), 0_1.

Burns, M. (2005). 3 lessons by Marilyn Burns: using storybooks to teach math. *Instructor, 114*(7), 27–30.

Caponera, E., Sestito, P., & Russo, P. M. (2016). The influence of reading literacy on mathematics and science achievement. *The Journal of Educational Research, 109*(2), 197–204.

D'Amore B. (2014). Insegnamento/Apprendimento significativo della matematica nella scuola dell'infanzia [Meaningful teaching/learning of mathematics in kindergarten], Convegno nazionale "La didattica della matematica: strumenti per capire e per intervenire", Tricase (Lecce), 3–4-5 marzo. http://www.dm.unibo.it/rsddm/it/articoli/damore/820%20Seminario%20Atti%20Convegno%20Tricase.pdf Accessed 24/04/2017.

Douville, P., Pugalee, D. K., & Wallace, J. D. (2003). Examining instructional practices of elementary science teachers for mathematics and literacy integration. *School Science and Mathematics, 103*(8), 388–396.

Draper, R. J. (2002). School mathematics reform, constructivism, and literacy: a case for literacy instruction in the reform-oriented math classroom. *Journal of Adolescent & Adult Literacy, 45*(6), 520–529.

Draper, R. J., Smith, L. K., Hall, K. M., & Siebert, D. (2005). What's more important—literacy or content? Confronting the literacy-content dualism. *Action in Teacher Education, 27*(2), 12–21.

Driver, M. K., & Powell, S. R. (2015). Symbolic and nonsymbolic equivalence tasks: the influence of symbols on students with mathematics difficulty. *Learning Disabilities Research & Practice, 30*(3), 127–134.

Duke, N. K., & Pearson, P. D. (2002). Effective practices for developing reading comprehension. In A. E. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction* (3rd ed., pp. 205–242). Newark, DE: International Reading Association.

Education, Audiovisual and Culture Executive Agency (2010). Gender differences in educational outcomes. Study on the measures taken and the current situation in Europe. Brussels: Education, Audiovisual and Culture Executive Agency P9 Eurydice. Retrieved from http://eacea.ec.europa.eu/education/eurydice/documents/thematic_reports/120EN.pdf.

European Commission (2010). A strategy for smart, sustainable and inclusive growth. http://ec.europa.eu/eu2020/pdf/COMPLET%20EN%20BARROSO%20%20%20007%20-%20Europe%202020%20-%20EN%20version.pdf. Accessed 02 Mar 2017.

European Commission. (2012). *EU high level group of experts on literacy. Final report*. Luxembourg: Publications Office of the European Union.

Grek, S. (2009). Governing by numbers: the PISA 'effect'in Europe. *Journal of Education Policy, 24*(1), 23–37.

Grimm, K. J. (2008). Longitudinal associations between reading and mathematics achievement. *Developmental Neuropsychology, 33*(3), 410–426.

Helwig, R., Rozek-Tedesco, M. A., Tindal, G., Heath, B., & Almond, P. J. (1999). Reading as an access to mathematics problem solving on multiple-choice tests for sixth-grade students. *The Journal of Educational Research, 93*(2), 113–125.

Hickendorff, M. (2013). The language factor in elementary mathematics assessments: computational skills and applied problem solving in a multidimensional IRT framework. *Applied Measurement in Education, 26*(4), 253–278.

INVALSI (2016). Rilevazioni nazionali degli apprendimenti 2015–16. Rilevazione degli apprendimenti nelle classi II e V primaria, nella classe III (Prova nazionale) della scuola secondaria di primo grado e nella II classe della scuola secondaria di secondo grado. [National learning assessment 2015–16. Second and fifth grades of primary school, third grade of lower secondary school, and second grade of upper secondary school.] https://invalsi-areaprove.cineca.it/docs/file/08_Rapporto_Prove_INVALSI_2016.pdf. Accessed 24/04/2017.

Jiban, C. L., & Deno, S. L. (2007). Using math and reading curriculum-based measurements to predict state mathematics test performance: are simple one-minute measures technically adequate? *Assessment for Effective Intervention, 32*(2), 78–89.

Johnson, B. E., & Zabrucky, K. M. (2011). Improving middle and high school students' comprehension of science texts. *International Electronic Journal of Elementary Education, 4*(1), 19.

Kane, R. B., Byrne, M. A., & Hater, M. A. (1974). *Helping children read mathematics*. New York: American Book Company.

Korhonen, J., Linnanmäki, K., & Aunio, P. (2012). Language and mathematical performance: a comparison of lower secondary school students with different level of mathematical skills. *Scandinavian Journal of Educational Research, 56*(3), 333–344.

Lafontaine, D., & Monseur, C. (2009). Gender gap in comparative studies of reading comprehension: to what extent do the test characteristics make a difference? *European Educational Research Journal, 8*(1), 69–79.

Liu, O. L., & Wilson, M. (2009). Gender differences in large-scale math assessments: PISA trend 2000 and 2003. *Applied Measurement in Education, 22*(2), 164–184.

Marks, G. N. (2008). Accounting for the gender gaps in student performance in reading and mathematics: evidence from 31 countries. *Oxford Review of Education, 34*(1), 89–109.

Marks, G. N., Cresswell, J., & Ainley, J. (2006). Explaining socioeconomic inequalities in student achievement: the role of home and school factors. *Educational Research and Evaluation, 12*(02), 105–128.

Marsh, H., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development, 76*(2), 397–416.

Martin, O., & Mullis, I. V. S. (Eds.). (2013). Relationships among reading, mathematics and science achievement at the fourth grade—implications for early learning. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College; Amsterdam: International Association for the Evaluation of Educational Achievement (IEA).

McConney, A., & Perry, L. B. (2010). Socioeconomic status, self-efficacy, and mathematics achievement in Australia: a secondary analysis. *Educational Research for Policy and Practice, 9*(2), 77–91.

MIUR (Ministero dell'Istruzione, dell'Università e della Ricerca) - Servizio Statistico. (2017). Focus "Gli immatricolati nell'a.a. 2016/2017 il passaggio dalla scuola all'università dei diplomati nel 2016." [Enrollment in the academic year 2016/2017: the transition from school to university of high school graduates in 2016.] http://www.miur.gov.it/documents/20182/639222/FOCUS+2016-17_immatricolazioni.pdf/3a1f576c-5498-4942-962c-435a2b791a5e?version=1.0 Accessed 17 Feb 2018.

Moyer, P. S. (2000). Communicating mathematically: children's literature as a natural connection. *The Reading Teacher, 54*(3), 246–255.

Mullis, I. V., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., & Preuschoff, C. (2009). *TIMSS 2011 assessment frameworks*. Amsterdam: International Association for the Evaluation of Educational Achievement.

Mullis, I.V.S. Martin MO., & Foy, P. (2013). The impact of reading ability on TIMSS mathematics and science achievement at the fourth grade: an analysis by item reading demands. In Martin, M.O and Mullis I.V.S. (Eds.), Relationships among reading, mathematics and science achievement at the fourth grade—implications for early learning. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College; Amsterdam: International Association for the Evaluation of Educational Achievement (IEA).

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). TIMSS 2015 international results in mathematics. TIMSS & PIRLS International Study Center at Boston College. http://timssandpirls.bc.edu/timss2015/international-results/ Accessed 24/04/2017.

Nasser, F., & Birenbaum, M. (2005). Modeling mathematics achievement of Jewish and Arab eighth graders in Israel: the effects of learner-related variables. *Educational Research and Evaluation, 11*(3), 277–302.

National Council of Teachers of Mathematics (Ed.). (2000). Principles and standards for school mathematics (Vol. 1). Reston, VA: National Council of Teachers of Mathematics.

Neufeld, P. (2006). Comprehension instruction in content area classes. *The Reading Teacher, 59*(4), 302–312.

OECD. (2010). *PISA 2009 results: what students know and can do—student performance in reading, mathematics and science* (Vol. I). Paris: OECD.
OECD. (2013). *PISA 2012 assessment and analytical framework: mathematics, reading, science, problem solving and financial literacy*. Paris: OECD.
OECD. (2014a). *PISA 2012 technical report*. Paris: OECD.
OECD (2014b). Proficiency scale construction. In OECD (Ed.), *PISA 2012 technical report* (pp. 287–303). Paris: OECD.
OECD. (2016). *PISA 2015 results (volume I): excellence and equity in education*. Paris: OECD.
Österholm, M. (2006). Characterizing reading comprehension of mathematical texts. *Educational Studies in Mathematics, 63*(3), 325–346.
Pica, T. (2002). Subject-matter content: how does it assist the interactional and linguistic needs of classroom language learners? *The Modern Language Journal, 86*(1), 1–19.
Pozio, S. (2011). La risoluzione di prove di competenza matematica. Analisi dei risultati italiani nell'indagine OCSE (Pisa, 2003) [Solution of mathematics competence tasks. Analysis of the Italian results in the OECD PISA 2003]. Roma: Nuova Cultura.
Reilly, E. M. (2014). Improving girls' interest in mathematics through writing. *Language, literacy, and learning in STEM education: research methods and perspectives from applied linguistics, 1*, 143.
Schleppegrell, M. J. (2007). The linguistic challenges of mathematics teaching and learning: a research review. *Reading & Writing Quarterly, 23*(2), 139–159.
Schmidt, W. H., Burroughs, N. A., Zoido, P., & Houang, R. T. (2015). The role of schooling in perpetuating educational inequality: an international perspective. *Educational Researcher, 44*(7), 371–386.
Siebert, D., & Draper, R. J. (2008). Why content-area literacy messages do not speak to mathematics teachers: a critical content analysis. *Literacy Research and Instruction, 47*(4), 229–245.
Snow, C. (2002). Reading for understanding: toward an R&D program in reading comprehension. Rand Corporation.
Stoet, G., & Geary, D. C. (2013). Sex differences in mathematics and reading achievement are inversely related: within-and across-nation assessment of 10 years of PISA data. *PLoS One, 8*(3), e57988.
Thurber, R. S., Shinn, M. R., & Smolkowski, K. (2002). What is measured in mathematics tests? Construct validity of curriculum-based mathematics measures. *School Psychology Review, 31*(4), 498.
Wilder, G. Z., & Powell, K. (1989). Sex differences in test performance: a survey of the literature. *ETS Research Report Series, 1989*(1), i–50.
Yore, L. D., & Treagust, D. F. (2006). Current realities and future possibilities: language and science literacy—empowering research and informing instruction. *International Journal of Science Education, 28*(2–3), 291–314.
Yore, L. D., Pimm, D., & Tuan, H. L. (2007). The literacy component of mathematical and scientific literacy. *International Journal of Science and Mathematics Education, 5*(4), 559–589.

Anna Maria Ajello. INVALSI (National Institute for the Educational Evaluation of Instruction and Training), Via Ippolito Nievo 35, 00153, Rome, Italy. E-mail: presidente@invalsi.it. Web site: http://www.invalsi.it. Department of Psychology, Sapienza Università di Roma, Via dei Marsi, 78, 00185 Rome, Italy. E-mail: annamaria.ajello@uniroma1.it. Web site: http://www.psicologia1.uniroma1.it/

Current themes of research:

Key competence and education. Economic knowledge and education. Social reasoning and interaction in classroom.

Most relevant publications in the field of Psychology of Education:

Morselli, D., & Ajello, A. (2016). Assessing the sense of initiative and entrepreneurship in vocational students using the European qualification framework. *Education+ Training*, *58*(7/8), 797-814.

Berti, A. E., Ajello, A. M., Aprea, C., Castelli, I., Lombardi, E., Marchetti, A., ... & Valle, A. (2017). Adolescents' and young adults' naïve understandings of the economic crisis. *Europe's Journal of Psychology*, *13*(1), 143.

Lombardi, E. & Ajello, A.M. (2017). *Ricerche di Psicologia, Vol. 40 (1), 93-123* ISSN 0391-6081, ISSNe 1972-5620

Elisa Caponera. INVALSI (National Institute for the Educational Evaluation of Instruction and Training), Via Ippolito Nievo 35, 00153, Rome, Italy. E-mail: elisa.caponera@invalsi.it. Web site: http://www.invalsi.it

Current themes of research:

Parental involvement at school. Gender difference in mathematics achievement. School effectiveness. Equity of school system.

Most relevant publications in the field of Psychology of Education:

Russo, P. M., Biasi, V., Cipolli, C., Mallia, L., & Caponera, E. (2017). Sleep habits, circadian preference, and school performance in early adolescents. *Sleep Medicine*, *29*, 20-22.

Caponera, E., Sestito, P., & Russo, P. M. (2016). The influence of reading literacy on mathematics and science achievement. *The Journal of Educational Research*, *109* (2), 197-204.

Caponera, E., & Russo, P. M. (2014). Student characteristics and mathematics achievement in TIMSS. *Cadmo, 22* (2)*,* 93-105.

Caponera, E. Russo P.M. (2010). Performance in TIMSS mathematics test: the influence of self-concept and socio-economic background. *Cadmo*, *18* (2), pp. 91-103.

Laura Palmerio. Department of Psychology, Sapienza Università di Roma, Via dei Marsi, 78, 00185 Rome, Italy. E-mail: laura.palmerio@invalsi.it. Web site: http://www.invalsi.it

Current themes of research:

Gender differences in achievement. Equity in education. Relations between reading literacy and mathematics. Civic and citizenship education.

Most relevant publications in the field of Psychology of Education:

Alivernini, F., Losito, B., & Palmerio, L. (2010). Le differenze nelle prestazioni degli studenti italiani in PISA 2006 e l'equità del sistema scolastico italiano: due studi. In *INVALSI. PISA 2006. Approfondimenti tematici e metodologici;* Roma, Armando; pp. 59-84.

Caponera, E., Losito, B., Pietracci, R., & Palmerio, L. (2016). Le differenze tra scuole nei risultati in matematica: un'analisi multilevel in funzione dei processi cognitivi indagati in PISA. In Palmerio, L. (ed.). *OCSE PISA 2012. Contributi di approfondimento*. Milano, FrancoAngeli; pp. 58-73.