Mathematics educators are speaking about PISA, aren't they?



*Published Version:*
Mathematics educators are speaking about PISA, aren't they? / Chiara Giberti; Andrea Maffia. - In: TEACHING MATHEMATICS AND ITS APPLICATIONS. - ISSN 0268-3679. - ELETTRONICO. - 39:4(2020), pp. 266-280. [10.1093/teamat/hrz018]

*Availability:*
This version is available at: https://hdl.handle.net/11585/894498 since: 2022-09-22

*Published:*
DOI: http://doi.org/10.1093/teamat/hrz018

# Mathematics educators are speaking about PISA, aren't they?

Chiara Giberti[1,*] and Andrea Maffia[2]
[1,*] Faculty of Education, Free University of Bozen, Bozen, Italy
For correspondence: chiara.giberti@unibz.it
[2] Department of Mathematics "F. Casorati", University of Pavia, Pavia, Italy

**Abstract :**
This paper sets out to explore the different uses made of OECD-PISA tests and data in mathematics education research. Through a comprehensive literature review of journals and conference papers, we show that although a large variety of topics is addressed, they do not cover all the topics considered in mathematics education research. Analysing the temporal and geographical distribution of papers, we find that there is increasing interest in the use of PISA in our field of research, and that different countries are involved in different ways in mathematics education research about PISA. As a conclusion, we suggest that critical research into the effect of PISA can be developed further, especially in those countries that have joined the OECD survey in recent years. Other future research paths using data from PISA are detected.

**Keywords:** PISA, international comparative studies, standardized assessment, survey, literature review.

## 1. Introduction

Researchers in education have taken an interest in the OECD (Organisation for Economic Co-operation and Development) Program for International Student Assessment (PISA) since its launch in 2000. We can find references to such tests in the literature of the very first years of the twenty-first century. Among these works, there is the discussion document of the fourteenth study of the ICMI (International Commission on Mathematical Instruction):

> Following the 2001 publication of results of the first PISA cycle (from 2000), an intense discussion has started, in several countries, about the aims and the design of mathematics instruction for schools, and especially about the role of mathematical modelling, applications of mathematics and relations to the real world. (Blum, 2002, p. 151)

However, mathematics' education researchers began to focus on PISA mainly after 2003. In fact, the main topic of the OECD test that year was mathematics, and many countries participated in PISA and in the Trends in International Mathematics and Science Study (TIMSS) promoted by IEA (International Association for the Evaluation of Educational Achievement). In the following years, PISA 2003 results (in grade 10) together with those of TIMSS 2003 (in grade 4 and 8) allowed a snapshot of mathematical learning in fourteen different countries distributed across all the continents. Ferrini-Mundy and Schmidt (2005) presented a longitudinal analysis of results from these international tests.

After the publication of results of these surveys, public interest in the learning of mathematics grew, at least in some countries (Sfard, 2005; Pons, 2012). However, research in mathematics education did not immediately adopt the results of international surveys for use. Sfard (2005) interviewed 74 important researchers in the field from all over the world, and reported that:

> Ed Silver […] marvels in one of his editorials that, "These days it seems that mathematics educators are a bunch of quantitatively competent individuals who are inclined to conduct qualitatively oriented studies." […] This qualitative preference of our respondents is counterbalanced by the increasingly popular international comparative studies, such as TIMSS and PISA, which focus mainly on students' measurable achievements. […] Our respondents do not

> help themselves to TIMSS or PISA findings even when responding to our questions on the state of mathematics education in their countries.

Our perception is that things have changed since 2005. Many publications show different modalities of use of PISA tests and related data, for the study of several topics. The innovation brought by the usage of international surveys was highlighted by ICMI, when the Freudenthal Medal was awarded to Frederick Koon Shing Leung. Indeed, ICMI justified the award by writing, among other things, that

> his ground-breaking work, for which he is internationally known, is the utilization of the perspective of the Confucian Heritage Culture to explain the superior mathematics achievement of East Asian students in international studies such as the IEA Trends in International Mathematics and Science Studies and the OECD Programme for International Student Assessment[1]

The embrace of PISA in the educational field has been counterbalanced more recently by a more critical approach to the effect of international surveys and multiple-choice tests (e.g. Große, 2017; Pons, 2012; Tsai & Li, 2017). Concerning our field of interest, Kanes, Morgan and Tsatsaroni (2014) defined the set of texts and technologies produced by OECD as the "PISA mathematics regime". Indeed, they stress how "these texts and technologies construct ways of thinking about mathematics, mathematical knowledge and mathematics education, about students, teachers and educational processes, and about how these are related to social and economic activity" (ibidem, p. 146). According to these authors, there are many references to PISA and related resources within mathematics education research. Furthermore, they sustain that PISA is often cited uncritically except for a few examples. However, Kanes and colleagues do not provide statistical evidence for these claims.

We pose the question of how much research actually pays attention to PISA, and how the results from the OECD international survey are used within mathematics education. Is it true that there are many references to PISA and related resources within mathematics education research, as stated by Kanes and colleagues (2014)? Has it always been true? Has it changed during time?
In the same work Kanes and colleagues also notice that "references to PISA were much more prevalent in CERME than in PME, perhaps reflecting the relative lack of attention to PISA among mathematics education researchers in the US" (2014, p.149). Indeed, a question arises: Does the interest to PISA depend on the geographical context of the research?
The aim of this paper is to explore the variety of uses of OECD-PISA tests and data in mathematics education research across the world and over time. We wish to highlight how PISA is generally used, but also to detect less common usages and their potentialities as future research paths. To this end, we present a large survey of research papers from journals and conferences about mathematics education.

## 2. Framework

A classification of papers was proposed by Owens (2013) in her survey about educational articles on PISA. In her work, she analyses 74 English papers from peer-reviewed journals. She considers journals about educational policies in the period 2000-2010. The first column of Table 1 summarizes Owen's classification of papers according to their topic.

Even though Owen does not refer specifically to the field of mathematics education, we have found that her categories are suitable for our purpose. In a previous work (Maffia & Giberti, 2016) we only selected papers from mathematics education journals (not from conference proceedings) published before 2014 and used her framework to classify them.
In the next paragraphs of this section we present the categorization derived that pilot analysis (ibidem).

---

[1] Retrieved from https://www.mathunion.org/icmi/awards/past-receipients/2013-hans-freudenthal-award [consulted 04/11/2019]

We found papers dealing with students' beliefs, attitudes, and identity and we classified them in the category *Affect and motivation*. Another Owen's category we used, is *Gender*: we included in this category all the papers having as main topic gender differences in learning outcomes.
We split Owen's category "National overview" into two sub-categories: *Mathematics education in a country* and *Comparative studies at national level*. In the first category we include those papers in which OECD-PISA data are used to comment about learning outcomes at national level or about national projects that were triggered by PISA outcomes.
When PISA data are used to compare differences between outcomes in two or more countries, we classify the paper in the category *Comparative studies at national level*.
In our classification (Maffia & Giberti, 2016) we decided to rename the category referring to teachers as *Teacher education* because all the papers we found about teachers concerned their initial training or professional development.

We discovered that some of Owen's categories did not apply to our sample of papers; hence, not all her categories were used.

We added three categories: *Comparative studies on tests*, *Curriculum development*, and *Modelling*. These categories emerge from articles concerning similar topics that we decided to regroup together.
In particular, *Comparative studies on tests* are those in which the main object of research is a test (usually at national level) that is comparable to PISA in terms of framework, typology of questions or results. Papers about *Curriculum development* investigate the process of modernisation of a national curriculum. Sometimes the PISA framework is used as a reference.
*Modelling* appears as the main mathematic topic studied through PISA data. This category can be considered as a sub-category of "Learning outcomes" in Owen's classification, but it is specific for mathematics.

We realised also that many papers focus on the PISA framework, while other works are more interested in the effect of the test on socio-cultural and political contexts. Moreover, such an extensive reading of the selected papers has revealed a second aspect: several publications that refer to OECD-PISA tests do not use the results of the tests or data belonging to the students' background questionnaire at all. Hence, we decided to carry out a second classification of the papers based on how OECD-PISA survey is used in the different papers (Maffia & Giberti, 2016). Since we did not find literature regarding this kind of classification, we adopted a ground-based approach to determine categories from our data (ibidem). First, we observed that researchers not only use data collected in the surveys but also use all the documents provided by OECD regarding PISA tests. In analysing the selected papers, we noticed the use both of documents provided before administration of the test – such as the theoretical framework – and documents released afterwards – such as the published test items. For this reason, we generated our categories based on the kind of document related to OECD-PISA surveys used in papers. According to the chronological order of the publication of the OECD-PISA documents, the first category includes those papers that make *Use of the theoretical framework*. In these works, the subject of the research is the framework itself; the aim is usually to revise some aspects or to explore the possibilities of applying it in other contexts. Furthermore, we observed that one of the most cited and discussed constructs is that of "Mathematical literacy" (Haara et al., 2017), which is often mentioned to make a comparison with similar definitions given in the context of other surveys, both at national and international levels.

In the second group of papers, categorised *Use of tasks*, authors analyse publicly disclosed PISA items. This happens generally in two different ways: texts of PISA-problems are used (with any modifications) in order to create problem-solving teaching sequences addressed to students, or PISA tasks are used to compose new tests targeted at students of different grades than the level investigated by the PISA survey, or even addressed to teachers.
Finally, in the category *Use of data* there are papers using data produced after the OECD tests were administered. We observe that the data used belong both to the mathematics test and to the students' background questionnaires, correlated where possible with the mathematics test results.
These three categories did not appear exhaustive in describing how OECD-PISA is used in mathematics education research; we observed that a small number of papers does not fit into this

classification. In fact, these papers, even though dealing with OECD surveys, do not use materials produced by OECD but are concerned with the analysis of effects of PISA tests on political, scientific, economic and social contexts.

**Table 1.** Comparison between classification by topic as developed by Owens (2013), by Maffia and Giberti (2016), and in the present work.

| **Owens' (2013) classification** | **Maffia & Giberti's (2016) classification** | **Final classification in this work** |
|---|---|---|
| Affect and Motivation | Affect and Motivation | Affect and Motivation |
| Gender | Gender | Gender |
| National overview | Mathematics education in a country | Mathematics education in a country |
| | Comparative studies at national level | Comparative studies at national level |
| Teacher and Classroom | Teacher education | Teacher education |
| | Comparative studies on tests | Comparative studies on tests |
| | Curriculum development | Curriculum development |
| | Modelling | Modelling |
| Technologies | | Technologies |
| Socio-economical background | | Equity |
| Immigration | | |
| | | Language |
| | | Textbooks |
| | | Lifelong education |
| Learning outcomes | | |
| Autonomy and management | | |
| Tracking | | |
| School choice | | |
| Retention | | |
| School entry age | | |
| Accountability | | |
| Efficiency | | |

## 3. Methods

The analysed sample of papers is taken from sources: *international journals* and *conference proceedings* on mathematics education. We selected journals according to the ratings of international journals proposed by the European society for Research in Mathematics Education (Törner & Arzarello, 2013). We selected those journals classified as level A or more. Among those journals, we analysed those which have an electronic research tool or that are inserted in an electronic database. The analysed papers come from the following journals: Educational Studies in Mathematics (ESM), ZDM-International journal on mathematics education, The Journal of Mathematical Behaviour (TJMB), Journal of Mathematics Teacher Education (JMTE), Journal of Research in Mathematics Education (JRME), and Mathematical Thinking and Learning (MTL).

We also considered papers from proceedings of international conferences. We selected large conferences (in terms of participant numbers) that are held regularly and on different continents: Annual conference of the International Group for the Psychology of Mathematics Education (PME), Annual conference of the North-American chapter of PME (PME-NA), Congress of European Research in Mathematics Education (CERME), Annual Conferences of the Mathematics Education Research Group of Australasia (MERGA).

At the moment of writing, the proceedings of these conferences as far as 2016 are all available to the authors. As PISA started in 2000, we decided to analyse all the papers published in the period 2001-2016. First, we used the databases search-tool (or the computer search-tool for PDF files of conferences proceedings) to look for all those papers in which the words 'PISA' or 'OECD' appear. Both authors read the abstract of these papers to select those that were pertinent (for instance, we excluded those papers in which the word 'Pisa' indicated just the authors' hometown). The year of publication of each paper was registered in order to analyse the temporal distribution.

The selection of papers from journals comprises 169 papers. Most of the papers belong to ZDM (64%). Nearly 20% belong to Educational Studies in Mathematics and 7% to Journal of Research in Mathematics Education. The remaining 10% is almost equally distributed between The Journal of Mathematical Behaviour, Journal of Mathematics Teacher Education and Mathematics Teaching and Learning.
We selected 351 papers belonging to proceedings of the main international conferences. The distribution of the papers is unequal: 44% of the papers belong to PME (56% if we consider the joint conferences with PMENA). There is a consistent presence of PISA-based studies in the Annual Conferences of the Mathematics Education Research Group of Australasia (15%) and in the Congress of the European society for Research in Mathematics Education (22%).

**Table 2.** Distribution of the papers selected from journals and from proceedings. We designate as 'PMENA-PME' the proceedings of those years in which the two conferences were held together.

| **Journal papers** | | **Conference papers** | |
|---|---|---|---|
| ZDM | 108 | PME | 154 |
| ESM | 32 | CERME | 77 |
| JRME | 12 | MERGA | 52 |
| TJMB | 7 | PMENA-PME | 42 |
| JMTE | 5 | PMENA | 26 |
| MTL | 5 | | |
| Total | 169 | Total | 351 |

A first coding was realised by topic: we used the same categories as in Maffia and Giberti (2016) – described in the previous section and listed in the second column of table 1. With the aim of collecting data about the geographical distribution, we registered information about the countries involved in those papers belonging to the topic-categories *Mathematics education in a country* and *Comparative studies at a national level*. We used only these two categories because all the papers belonging to them state explicitly the countries studied. PISA list of participating countries was used for this classification[2].

After categorizing all the papers according to the framework proposed in the previous section, we realized that almost 19% of journal and conference papers did not belong to any category. We returned to those papers to search for the topics they addressed. First, we used Owens' (2013) categories and we found that there were papers about *Technology*, *Socio-Economical background*, and *Immigration*. Hence, we included these categories in the classification: papers dealing with the

[2] According to OECD-PISA classification we used *China* to refer to the four PISA-participating Chinese provinces: Beijing, Shanghai, Jiangsu and Guangdong. Hong Kong, Macao and Chinese Taipei are considered separately.

introduction of new technologies in teaching or assessing were included in the former category; since all the papers about social context or immigration concerned different opportunities in learning, we grouped the two Owen's categories labelling them as *Equity*. Then we adopted a grounded approach (Cohen, Manion, & Morrison, 2007) to identify new categories for the remaining papers. From an open coding procedure, three categories emerged. In some papers, results in PISA mathematics test and home language test are compared; we coded these papers as *Language*. Those articles having textbooks as their unit of analysis were inserted in the category *Textbooks*. Finally, we found papers about mathematics education outside schools, for adults and/or along the lifespan. We labelled them as *Lifelong education*. The complete list of the categories used in the present work is shown in the third column of table 1.

Our second main coding procedure was realized for all the papers according to their use of PISA documents. In this case, the three categories *Use of framework*, *Use of tasks*, and *Use of data* – presented in the previous section – were exhaustive. We highlight that, for both classifications, some papers may belong to more than one category; in such cases, we assigned the papers to all suitable categories.

The two authors analysed the papers individually using the categories described in this section. Comparing the classifications by the two authors, we found that 12% of cases had different classification; we discussed those cases with an experienced researcher in the field to decide which category was more suitable. In those cases in which we did not reach a consensus, the paper was classified as *Other*.

## 4. Data analysis

We stated that the aim of this paper is to explore the variety of uses of OECD-PISA survey in mathematics education research across the world and over time. For this reason, we will first present results about temporal and geographical distribution of papers in our sample. Furthermore, we posed a question about how PISA is generally used; adopting the presented framework we analyse our sample of papers both in terms of the topic addressed and the usage of PISA documents.

### 4.1. Temporal distribution

We observe an increasing interest in OECD-PISA tests: while till 2003 the number of papers using PISA was less than five both for conferences and journals, in recent years the usage of PISA in mathematics education is visibly more popular (fig. 1). However, there is not a strong correlation between the years and the number of papers published in journals (ρ=0.50). We observe that the situation is different for conferences; indeed, there is a strong correlation between the years and the number of published conference papers (ρ=0.87).

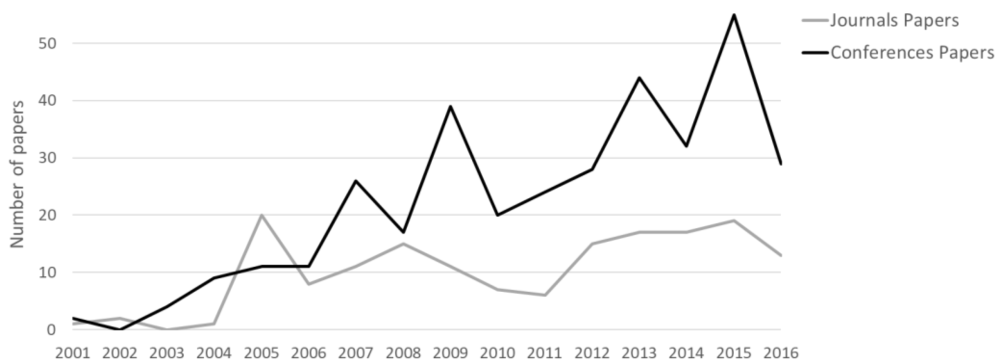


**Figure 1.** Time distribution of the selected papers in the journals and conferences.

While PISA surveys took place in 2000, 2003, 2006, 2009, 2012 and 2015, we observe that in 2005 and 2008 we have the largest number of journal papers regarding PISA in the first 10 years. The considerably large number of publications in 2005 and 2008 is due to research studies based on results of PISA tests administered in 2003 and 2006; the two years between the test and the publication may constitute the time required for the restitution of data and the publication of the papers. While the trend of the number of papers using PISA in journals fluctuates (we observe a real increase only in recent years), the same is not true of the distribution of papers in conferences. The time needed to write and present a research in a congress is usually less than the time it takes to publish a paper in a journal - indeed it could be possible to publish a paper using test data even just one year after the test administration.

### 4.2. Geographical distribution

For this analysis we considered only those papers belonging to the topic-categories *Mathematics education in a country* and *Comparative studies at a national level*, presented in detail in the next section. The considered sub-sample comprises 74 journal articles and 123 conference papers. We counted the number of papers explicitly involving one or more countries and drafted the maps below (figs. 2 and 3).


**Figure 2.** Geographical distribution of journal papers.


**Figure 3.** Geographical distribution of conference papers.

We can observe that the most mentioned country is Germany (17 journal papers and 13 conference papers). However, we can also find a high number of papers about eastern countries like Korea (journals: 15, conferences: 9), Taiwan (journals: 2, conferences: 13), China (journals: 6, conferences: 5), and Japan (journals: 5, conferences: 5). Usually, they are compared with western countries in Europe or North America; indeed, among the papers classified as *Comparative studies at a national level*, 26 (journals: 14, conferences: 12) compare at least an eastern country to one or more western countries. As Owens (2013) already noticed in the field of general education, also in mathematics education there is an imbalance in the study of PISA results around the world. Europe, North America and Australia are the areas to which most of the research on PISA refers. Asia and South America are mentioned in fewer studies, and Africa is almost not represented at all. It is also true that the African continent had fewer countries (yet still some, like Algeria and Tunisia) which participated in the PISA survey.

### 4.3. Topic classification

We analysed the selected papers, separating the papers belonging to journals from those belonging to conferences. For both groups we constructed a classification by topic and a classification by usage, as described in the theoretical framework. As mentioned above, the classification by topic is based on the categories that we derived from Owens' classification (Owens, 2013) and adapted for our study. Using the 8 categories, we were able to classify nearly 80% of papers belonging both to journals and conferences.

Table 3 reports the number of papers published in the selected journals and proceedings for each topic. The total number is greater than the total number of selected papers because, when we found papers about two topics, we opted to include them in both topics.

**Table 3.** Classification by topic of papers published in journals (second column) and in conferences (third column). Some papers are included in more than one topic.

| Topic | Number of papers in journals | Number of papers in conferences |
|---|---|---|
| Mathematics education in a country | 51 (30%) | 87 (25%) |
| Comparative studies at national | 23 (14%) | 36 (10%) |

| | | |
|---|---|---|
| level | | |
| Teacher education | 21 (12%) | 80 (23%) |
| Comparative studies on tests | 17 (10%) | 20 (6%) |
| Curriculum development | 16 (10%) | 31 (9%) |
| Gender | 10 (6%) | 23 (7%) |
| Affect and Motivation | 5 (3%) | 38 (11%) |
| Modelling | 7 (4%) | 49 (14%) |
| Technology | 7 (4%) | 10 (3%) |
| Equity | 8 (5%) | 11 (3%) |
| Language | 1 (1%) | 10 (3%) |
| Textbooks | 3 (2%) | 7 (2%) |
| Lifelong Education | 2 (1%) | 6 (2%) |
| Other | 24 (14%) | 46 (13%) |
| Total | 169 | 351 |

A large part of the papers (around 30% for journals and 25% for conferences) addresses the issue of *Mathematics education in a country*. In these studies, OECD-PISA data are often cited as a trigger for new works and investigations and only in some cases are the main focus of the research. A paradigmatic example of popular use of the OECD-PISA survey is provided by papers from Germany, a country where international surveys such as PISA are flanked by national and regional standardised tests, and the importance given to these surveys and related research has ensured also important changes in mathematics curricula (Lorenz, 2005; Neubrand et al., 2001). We found that many works published both in journals (ZDM in particular) and in conference proceedings, are focused on mathematics education in Germany (see also the section about Geographical distribution) and, in this context, they also analyse in-depth differences between the 16 federal states and between the different school systems.

*Comparative studies at national level* often start by observing strong differences in the results of the OECD-PISA tests in two or more countries, before going on to identify the causes in the characteristics of the educational systems. Many of these studies (journals: 14, conferences: 12) focused on comparison between western and eastern countries because, in standardised assessment, eastern countries usually achieve better results. A typical example of a paper on this issue is offered by Yang & Leung (2011) in which the authors highlight the importance of social and cultural factors for the development of mathematics education studies; indeed, the approaches to this discipline in the east and west are different. These first two categories (*Mathematics education in a country* and *Comparative studies at national level*) are awarded particular attention, especially in journal papers: these two topics constitute almost half of the selected papers (confirming the results of our previous study Maffia & Giberti, 2016). Concerning conference papers, there are also other topics with many studies but, still, these two categories cover over a third of the selected works.

The third category refers to *Teacher education*, which consists in around 12% of journal papers and 23% of conference papers. We observe that OECD-PISA survey is used both to evaluate difficulties in mathematics of future teachers (Sàenz, 2009) and to experiment new training methods, for example based on the introduction of new technologies (Goos & Geiger, 2012).
The information gathered by the OECD survey is combined with those of other tests administered locally: several papers (around 10% of journal papers and 6% of conference papers) compare the features of the tests and the results of the OECD-PISA survey with similar tests administered at a national level. We include these papers in the category *Comparative studies on tests*, and we observe an interest in comparison with other standardised tests (e.g. Lorenz, 2005).

On a national level, the OECD-PISA survey and its framework contribute to the development of mathematics curricula, and about 10% of the analysed articles (from both journals and conference

papers) deal with this topic. We found that in many countries, such as Japan (e.g. Abe, 2007), the OECD-PISA framework had a great influence on national curricula.
The sixth identified category includes around 6% of journal papers and 7% of conference papers and concerns research on *Gender* differences. Interest in this issue, which is on a general level and focused on a single nation, has increased in recent years, also as a result of the developing role of OECD-PISA surveys in research in mathematics education (Leder & Forgasz, 2008). A paradigmatic example of the use of OECD-PISA surveys to analyse different mathematical performances by males and females is the work by Steinthorsdottir and Sriraman (2008, p. 591); the authors analyse the results of students in Iceland and in other countries. They observe that "Iceland was the only country in which the mathematics gender gap favoured girls".

The category *Affect and Motivation* is much more common in conference papers (11%) than journal papers (3%). We found several studies investigating the role of affective and metacognitive factors in mathematics teaching/learning processes and, using OECD-PISA results, they show how students' performance in mathematics strongly depends on the context (Roesken et al., 2011).

The role of *Technologies* in mathematics education is highly studied (in 3% of the analysed conference papers and 4% of journal papers). Some researchers focus their attention on the use of technologies as assessment tools. Jacinto, Amado and Carreira (2009) use the cognitive processes as listed in the PISA 2003 framework to analyse the outcomes of mathematical competition on the internet. Papers in the category *Equity* (5% of journal papers and 3% of conference papers) deal with the issue of equity in the school system, meaning the difficulty that some students face (because of socio-cultural background) in achieving the same results as others. For instance, Zevenbergen, Niesche, Grootenboer and Boaler (2008) notice that Australia performed well on international comparisons in terms of overall performance in mathematics, but it was one of the poorest-performing countries in terms of equity.

We found also other themes emerging from the remaining papers. We counted that, for conference papers, 3% of papers concern *Language* and just one journal paper fits into this category. Performances in the PISA survey depends strongly on students' understanding of the language in which the test is written. Researchers pay attention to the relation between mathematics achievement and home-language results. An example is given by Bergqvist, Theens and Österholm (2016), who found that there is no correlation between students' results in mathematics and the length of sentences or words in three different languages.
Another topic tackled by researchers in mathematics education is the analysis of *Textbooks* and 2% of papers belong to this category both in journal and conference papers. A paradigmatic example of this category is provided by Bispo and Ramalho (2005). They use the PISA framework as a categorisation tool for tasks in 9th grade high school textbooks.
The final category, *Lifelong education* is more common in conference papers (2%). A paradigmatic example is provided by Gal (2013) who studies adult learning while considering mathematics skills investigated by PISA surveys and by the Programme for International Assessment of Adult Competencies (PIAAC) both promoted by the OECD.
Finally, considering also these new topics, we categorised almost 90% of the papers from journals and conferences. We inserted into the category *Other* all the papers that did not fit in the categories mentioned above.

**4.4. Usage classification**

Many of the articles mentioning the OECD PISA survey do not use directly data from the mathematics test, but use other information gathered by OECD PISA surveys. We therefore form a classification of the papers based on the way in which the PISA surveys are used in mathematics education research. We use the categories that we identified in a previous work (Maffia & Giberti, 2016) and that are presented above in section 3.

The first group identified includes the papers that make *use of the framework*. For instance, Venkat and Winter (2015) compare how mathematical literacy is conceived in PISA with the definition that is given in South-African national standards. We found that the usage of the PISA framework is often present and discussed also in papers belonging to conferences. For instance, in his plenary lecture at CERME8, Boero (2013, p. 26) criticizes the definition of mathematical literacy because it "may convey an image of mathematics as an absolute and homogeneous body of knowledge and tools to solve every kind of problem".

The second group of papers makes *use of the tasks* publicly available. This generally happens in two different ways: the texts of the proposed problems are used (perhaps with modifications) to produce problem-solving teaching sequences addressed to the students (e.g. O'Shea & Leavy, 2013) or they are assembled to compose new tests. In many papers, one or more questions are administered to students at school levels different from those of the PISA survey. An example is given by the work by Lin and Ufer (2016), who used eye-tracking to check differences between Taiwanese and German students' behaviours in solving three PISA tasks. There are also works in which the PISA questions are administered to teachers (e.g. Sàenz, 2009).

The third category of research involves the *use of the data* that are returned after the administration of the OECD test. The used data may be either the results of the mathematics test and the answers to the student questionnaire, possibly related to the mathematics test (Heinze, Reiss & Franziska, 2005). The "use of the data" is the most widespread use. Often, it is a very superficial use of data: researchers claim that one country had a better result than another, or they simply state that students from a certain country had difficulties in some topic. Hence, our results are in line with the conclusion reached by Kanes et al. (2014): PISA results are often used in a superficial way. But this is not always the case. For instance, Forgasz (2008) compares gender differences in PISA outcomes to show that female disadvantage increased from 2003 to 2006 in Australia.

Although these categories may seem exhaustive, it emerged that a small percentage of items did not fit into any of the three. These publications do not deal directly with evidence produced by the OECD surveys, nor do they make use of material produced by PISA, but rather they refer to the effects on the scientific, political, economic or social contexts. An example of this type of article can be identified in the work of Leder and Forgasz (2008) who, working on gender differences, note that "The results of large scale international testing [...] have attracted widespread attention from the general mathematics education research community as well as from those with a particular interest in gender differences in mathematics learning" (p. 518).

Using this classification, we observe that most of the papers use the data and results of the PISA surveys: this category covers around 60% for both journals (102 papers) and conferences (211 papers). Moreover, use of the PISA framework is also widespread, emerging in around 30% of papers (51 journal papers and 124 conference papers).
Use of questions appears in 17 journal papers (10%) and in 45 conference papers (13%). The study of PISA effects is the focus of 12 journal papers (7%) and 23 conference papers (7%). We point out that the overall sum is over 100% because some papers belong to more than one category.

**Table 4.** Classification by usage of papers published in journals (second column) and in conferences (third column). Some papers are included in more than one usage.

| Usage | Number of papers in journals | Number of papers in conferences |
|---|---|---|
| Use of the framework | 51 (30%) | 124 (35%) |
| Use of the tasks | 17 (10%) | 45 (13%) |
| Use of the data | 102 (60%) | 211 (60%) |
| Other | 12 (7%) | 23 (7%) |
| Total | 169 | 351 |

## 5. Discussion and Conclusion

We posed the question of how much research actually pays attention to PISA, and how the results from the OECD international survey are used within mathematics education. Our study shows that there are many references to PISA and related resources both in mathematics education journals and conferences, even if not equally distributed. Among journals, ZDM has a higher percentage of papers about PISA. The general number of papers (about PISA or not) that are annually published by ZDM is not greater than the other journals, so we can state that the issue of PISA tests is not represented equally among the considered journals about mathematics education. This claim is certainly true in our sample, but we must highlight that we could only consider six journals because of our methodological choices.
Kanes and colleagues notice that "references to PISA were much more prevalent in CERME than in PME, perhaps reflecting the relative lack of attention to PISA among mathematics education researchers in the US" (2014, p.149). Our results refute the claim that reference to PISA are more prevalent in CERME than in PME, even considering that CERME is biannual. We can confirm the statement about the US: the papers presented in PME-NA were less then those presented in CERME even if the latter conference takes place less frequently. However, the US results are often compared with those of other countries in published journal articles, as shown in figure 2. A similar observation is possible for Eastern-Asian countries that are often considered in international comparison even if there are not many references to PISA in MERGA. This last consideration might depend on the fact that MERGA takes often place in Australia and we did not consider any conference taking regularly place in Asia.

Considering the question about the change of interest toward PISA during time, we have shown that the number of papers in journals fluctuates while it increases in conferences. It would be interesting to develop the research to understand the reasons for this difference. Furthermore, we stress that the different countries are involved differently in research. Western and Central Asia, South America and Africa are still underrepresented in mathematics education research using PISA. There is still the possibility to deepen research at national level in some countries, especially in the new countries involved in PISA 2018 and in those that will join PISA in future surveys.

Researching how PISA is used in mathematics education, first we classified the papers according to the addressed topic. Adapting Owens' (2013) classification, we developed and tested a new classification which is specific to the field of mathematics education. This classification highlighted different distribution among journals and conferences: for instance, there are still few research studies in mathematics education journals using data from PISA items on affect and motivation towards mathematics. There are more conference papers on this topic.
On the other hand, the categorisation by usage leads to similar results both for conferences and journals. The *use of data* is widespread but, in accordance with previous research (Kanes et al., 2014), it is often (but not always) very superficial. We can see that the use of questions from PISA still constitutes a small category. We also observed that critical research on the effects of PISA is still limited and could be developed. This appears as particularly relevant and we believe that those countries that were involved for the first time in PISA during the last test session may offer important starting points for research into the effect of PISA on the public opinion of mathematics and on curricular resources. Another possible conjecture is that there is an effect of the PISA view of mathematics also on non-participant countries.

It is interesting to notice that, by intersecting the categorisations, we can obtain some additional information. Table 5 presents the intersection between classifications by topic and by use.

**Table 5.** Cross-classification of papers.

| | **Use of data** | | **Use of tasks** | | **Use of framework** | | **Other** | |
|---|---|---|---|---|---|---|---|---|
| | Journ. | Conf. | Journ. | Conf. | Journ. | Conf. | Journ. | Conf. |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Maths education in a country | 41 | 67 | 3 | 4 | 11 | 23 | 3 | 7 |
| Comp. studies at national lev. | 22 | 29 | 0 | 3 | 1 | 8 | 0 | 1 |
| Teacher education | 9 | 41 | 6 | 12 | 6 | 27 | 0 | 3 |
| Comparative studies on tests | 3 | 8 | 2 | 3 | 13 | 13 | 0 | 0 |
| Curriculum development | 5 | 15 | 2 | 3 | 9 | 16 | 3 | 0 |
| Gender | 9 | 22 | 0 | 1 | 0 | 2 | 1 | 0 |
| Affect and Motivation | 5 | 33 | 0 | 4 | 0 | 0 | 0 | 1 |
| Modelling | 1 | 11 | 0 | 9 | 6 | 26 | 0 | 4 |
| Technology | 4 | 3 | 0 | 1 | 3 | 4 | 0 | 2 |
| Equity | 8 | 10 | 1 | 1 | 1 | 1 | 0 | 1 |
| Language | 0 | 7 | 1 | 2 | 1 | 1 | 0 | 0 |
| Textbooks | 3 | 4 | 0 | 0 | 0 | 3 | 0 | 0 |
| Lifelong education | 0 | 1 | 1 | 1 | 2 | 4 | 0 | 0 |
| Other | 11 | 25 | 2 | 9 | 5 | 21 | 5 | 4 |

We can see that some combinations are frequently explored (like the use of results for studying mathematics education outcomes in a country), while others are represented by few or no papers. Some combinations could be impossible to achieve, but those that have even a low number of papers could represent new research opportunities. Hence, an important result of this work is that of evidencing possibilities for research in mathematics education making use of the OECD-PISA survey. For instance, there are no papers about the effects of PISA on textbooks, a topic that appears relevant if it is true that this test can affect the image of mathematics in school (Kanes et al., 2014; Boero, 2013). Furthermore, recently the administration of PISA is computer-based. This could have an effect not only on students' results but also on the way in which computers are involved in teaching mathematics. This short list offers just some examples of new routes of research that are still underrepresented in the literature, highlighted by the results of our analysis.

We stress that the one used in this work is not an exhaustive list of topics of research in mathematics education. Some Owen's (2013) categories were never used, but they could apply to mathematics education. For instance, the topic *School Choice* could be investigated in relation to PISA mathematics outcomes. Similar research might be carried out considering *School Entry Age*. Furthermore, mathematics education research addresses several other topics as pointed out by Inglis and Foster (2018). Comparing their list to ours we can observe that topics related to specific mathematical contents (e.g. *Addition and Substraction*, *Euclidean Geometry*) or abilities (e.g. *Spatial Reasoning*, *Proof and Argumentation*) are rarely represented or missing in research concerning PISA.

Further development of our research could include a classification according to the theoretical framework and/or typology of analysis in order to understand if new theoretical tools are emerging or can be used to study PISA in the context of mathematics education. Finally, the methods here presented could be applied also to studying the role of PISA (or other large-scale surveys) in science education (Elmas et al., 2018).

## References


Abe, Y. (2007). Perspective on Mathematical Literacy in Japan: Towards Curriculum Construction. *Proceedings of the 31st Conference of the International Group for the Psychology of Mathematics Education* (J.H. Woo, H.C. Lew, K.S.Park & D.Y. Seo eds), vol. 1. Seoul: IGPME, p. 301.

Bergqvist, E., Theens, F., and Österholm, M. (2016). Relations between linguistic features and difficulty of PISA tasks in different languages. *Proceedings of the 40th Conference of the International Group for the Psychology of Mathematics Education* (C. Csikos, A. Rausch & J. Szitanyi eds). Szeged, HU: IGPME, p. 125.

Bispo, R., and Ramalho, G. (2005). Evaluation of suggested problems in Portuguese mathematics textbooks. *Proceedings of the 29th annual conference of the International Group for the Psychology of Mathematics Education* (H. L. Chick & J. L. Vincent eds), vol. 1. Melbourne, AUS: IGPME, p. 304.

Blum, W. (2002). ICMI Study 14: Applications and modelling in mathematics education – Discussion document. *Educational Studies in Mathematics, 51*(1-2), 149-171.

Boero, P. (2013). Mathematics education today: scientific advancements and societal needs. *Proceedings of CERME8* (B. Ubuz, C. Haser & M.A. Mariotti eds.). Antalya, TR: ERME, pp. 8-32.

Elmas, R., Bodner, G. M., Aydogdu, B., and Saban, Y. (2018). The Inclusion of Science Process Skills in Multiple Choice Questions: Are We Getting Any Better?. *European Journal of Science and Mathematics Education, 6*(1), 13-23.

Ferrini-Mundy, J., and Schmidt W.H. (2005). International Comparative Studies in Mathematics Education: Opportunities for Collaboration and Challenges for Researchers. *Journal for Research in Mathematics Education, 36*(3), 164-175.

Forgasz, H. (2008). Gender, socio-economic status, and mathematics: Performance among high achievers. *Proceedings of the Joint Conference PME 32 and PME-NA 30* (O. Figueras, J. L. Cortina, S. Alatorre, T. Rojano & A. Sepúlveda eds). Morelia, MX: IGPME, pp. 25-32.

Gal, I. (2013). Mathematical skills beyond the school years: A view from adult skills surveys and adult learning. *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education* (A. M. Lindmeier & A. Heinze eds), vol. 1. Kiel, DE: IGPME, pp. 31-46.

Goos, M., and Geiger, V. (2012). Connecting social perspectives on mathematics teacher education in online environments. *ZDM – The International Journal on Mathematics Education, 44*(6), 705-715.

Große, C. S. (2017). Effects of Multiple Choice Options in Mathematics Learning. *European Journal of Science and Mathematics Education, 5*(2), 165-177.

Haara, F. O., Bolstad, O. H., and Jenssen, E. S. (2017). Research on Mathematical Literacy in Schools--Aim, Approach and Attention. *European Journal of Science and Mathematics Education, 5*(3), 285-313.

Heinze, A., Reiss, K., and Franziska, R. (2005). Mathematics achievement and interest in mathematics from a differential perspective. *ZDM – The International Journal on Mathematics Education, 37*(3), 212-220.

Inglis, Matthew and Foster, Colin (2018) Five decades of mathematics education research. *Journal for Research in Mathematics Education*, *49* (4). pp. 462-500. ISSN 1945-2306

Jacinto, H., Amado, N., and Carreira, S. (2009). Internet and mathematical activity within the frame of "Sub 14". *Proceedings of the Sixth Congress of the European Society for Research in Mathematics Education* (V. Durand-Guerrier, S. Soury-Lavergne & F. Arzarello eds*).* Lyon, FR: Institut National de Recherche Pédagogique and ERME, pp. 1221-1230.

Kanes, C., Morgan, C., and Tsatsaroni, A. (2014). The PISA mathematics regime: Knowledge structures and practices of the self. *Educational Studies in Mathematics, 87*(2), 145-165.

Leder, G., and Forgasz, H. (2008). Mathematics education: new perspectives on gender. *ZDM– The International Journal on Mathematics Education, 40*(4), 513-518.

Lin, T.W., and Ufer, S. (2016). German and Taiwanese students' eye movements when solving statistics items. *Proceedings of the 40th Conference of the International Group for the Psychology of Mathematics Education* (C. Csikos, A. Rausch & J. Szitanyi eds). Szeged, HU: IGPME, p. 195.

Lorenz, J. H. (2005). Zentrale Lernstandsmessung in der Primarstufe—Vergleichsarbeiten Klasse 4 (VERA) in sieben Bundesländern. *ZDM – The International Journal on Mathematics Education, 37*(4), 317-323.

Maffia, A., and Giberti, C. (2016). Didattica della matematica e PISA: strade percorse e nuovi sentieri da battere. *PISA 2012. Contributi di approfondimento* (L. Palmerio ed). Rome, IT: Franco Angeli, pp. 637-644.

Neubrand, M., Biehler, R., Blum, W., Cohors-Fresenborg, E., Flade, L., Knoche, N., Lind, D., Löding, W., Möller, G., and Wynands, A. (2001). Grundlagen der Ergänzung des internationalen PISA-Mathematik-Tests in der deutschen Zusatzerhebung. *ZDM – The International Journal on Mathematics Education, 33*(2), 45-59.

O'Shea, J., and Leavy, M.L. (2013). Teaching mathematical problem-solving from an emergent constructivist perspective: the experiences of Irish primary teachers. *Journal of Mathematics Teacher Education, 16*(4), 293-318.

Owens, T. L. (2013). Thinking beyond league tables: A review of key PISA research questions. PISA, Power and Policy: the emergence of global educational governance. *Oxford Studies in Comparative Education, 23*(1), 27-49.

Pons, X. (2012). Going beyond the 'PISA shock' discourse: an analysis of the cognitive reception of PISA in six European countries, 2001-2008. *European Educational Research Journal, 11*(2), 206-226.

Roesken, B., Pepin, B., and Toerner, G. (2011). Beliefs and beyond: affect and the teaching and learning of mathematics. *ZDM – The International Journal on Mathematics Education, 43*(4), 451-455.

Sàenz, C. (2009). The role of contextual, conceptual and procedural knowledge in activating mathematical competencies (PISA). *Educational Studies in Mathematics, 71*(2), 123-143.

Sfard, A. (2005). What Could be More Practical than Good Research?. *Educational Studies in Mathematics, 58*(3), 393-413.

Steinthorsdottir, O. B., and Sriraman, B. (2008). Exploring gender factors related to PISA 2003 results in Iceland: a youth interview study. *ZDM – The International Journal on Mathematics Education, 40*(4), 591-600.

Törner, G., and Arzarello, F. (2013). Grading mathematics education research journals. *Mitteilungen der Gesellschaft für Didaktik der Mathematik, 39*(95), 31-34.

Tsai, T. L., and Li, H. C. (2017). International comparative studies in mathematics education: are we obsessed with the international rankings of measured educational outcomes*?. International Journal of Mathematical Education in Science and Technology, 48*(8), 1262-1267.

Venkat, H., and Winter, M. (2015). Boundary objects and boundary crossing for numeracy teaching. *ZDM – The International Journal on Mathematics Education, 47*(4), 575-586.

Yang, X., and Leung, F. K. S. (2011). Mathematics teaching expertise development approaches and practices: similarities and differences between Western and Eastern countries. *ZDM – The International Journal on Mathematics Education, 43*(6-7), 1007-1015.

Zevenbergen, R., Niesche, R., Grootenboer, P., and Boaler, J. (2008). Creating equitable practice in diverse classrooms: Developing a tool to evaluate pedagogy. *Navigating currents and charting directions* (M. Goos, R. Brown & K. Makar eds), vol. 2. St Lucia, AU: MERGA, pp. 637-644.

**Authors' Biographies**

**Chiara Giberti** (corresponding author), Free University of Bozen, Italy.
Facoltà di Scienze della Formazione
viale Ratisbona, 16
39042 Bressanone (BZ)
Italy
E-mail: chiara.giberti@unibz.it

ORCID: 0000-0001-7446-6709

**Chiara Giberti** is Ph.D. in Math education and fellow at the Free University of Bozen. She collaborates with INVALSI and her research interest focuses on the potential of standardized assessment in mathematics education research. In particular, she studies gender differences in large scale math tests from both a statistical and a didactical point of view. She also teaches mathematics and science in middle school and she collaborates in teacher training courses in math education.

**Andrea Maffia**, Department of Mathematics “F. Casorati”, University of Pavia, Pavia, Italy
Via Ferrata, 5
27100 Pavia
Italy
E-mail: andrea.maffia@unipv.it.

**Andrea Maffia**, Ph.D. in Mathematics Education, works as researcher at the Department of Mathematics of the University of Pavia, Italy. His research interests concern the semiotic aspects involved in the process of mathematics teaching and learning, with particular reference to relational thinking in arithmetic and algebra. As representative of young researchers, he is a member of the board of the European Society for Research in Mathematics Education.