View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by The University of Sydney: Sydney eScholarship Journals online

International Journal of Innovation in Science and Mathematics Education, 29(1), 1-15, 2021

Quantifying Solid Findings in Mathematics Education: *Loss of Meaning* for Algebraic Symbols

Giorgio Bolondi[a] and Federica Ferretti[a]

Corresponding author: Giorgio Bolondi (giorgio.bolondi@unibz.it)
[a] Free University of Bozen-Bolzano, Faculty of Education, 39121 Brixen, Italy.

Keywords: algebraic modelling, large-scale assessment, solid findings, meaning of symbols

Abstract

We report an example of a research approach aimed at gathering quantitative pieces of evidence of solid findings in mathematics education. The main goal of this project is to provide an additional perspective on solid findings in education, to be used by teachers and by researchers in their work. As a case study, we present a situation of "loss of meaning" in algebra, exploring it with data coming from a large-scale assessment interpreted by means of theoretical lenses. We are able to give information about the extent of the phenomenon and to highlight how the phenomenon is relevant also for high-level students. This approach can provide a link between large-scale assessment results, educational research, and teachers' practices, and suggests further research issues.

Introduction and Overview

Overview of the problem

Mathematics education has its solid findings. As defined by European Mathematical Society (EMS, 2011), solid findings are findings that:

- result from trustworthy, disciplined inquiry, thus being sound and convincing in shedding light on the question(s) they set out to answer.
- are generally recognised as important contributions that have significantly influenced and/or may significantly influence the research field.
- can be applied to circumstances and/or domains beyond those involved in this particular research.
- can be summarised in a brief and comprehensible way to an interested but critical audience of non-specialists (especially mathematicians and mathematics teachers).

A first list of proposed "solid findings" was presented in a series of issues of the European Mathematical Society Newsletter: included, among others, Didactic Contract (EMS, 2012), Sociomathematical Norms (EMS, 2013), Linearity (EMS, 2015) and, on a topic which has intersections with our case study, Models and Modelling (Niss, 2012). An updating of the discussion on what should be considered a "solid finding" in mathematics education research is contained in Bosch, Dreyfus, Primi, and Shiel (2017) and Dreyfus (2019). Different perspectives are possible. Looking to quantitative data, a psychometric point of view on the "robustness" of a finding is proposed in a particular case by Primi (2017).

Results in mathematics education are obtained and validated through shared research paradigms and methodologies, with a prevalence of a qualitative approach (Hart, Smith, Swartz, and Smith, 2009). For many didactic phenomena, there are detailed descriptions and general theoretical frameworks, and these descriptions and frameworks are also useful tools for the teachers when they face what happens in a classroom and when they try to interpret their students' behaviour. What is sometimes missing is a quantification of the magnitude of the phenomenon, which might help both teachers in understanding their specific teaching-learning situations and researchers in highlighting the articulation and the extension of the phenomenon. Having results related to the magnitude and the spread of a phenomenon in a specific systemic context might be a tool for linking the findings of the research to the on-the-field practice of teachers.

This paper is the first step of a general research project, aimed at collecting robust quantitative pieces of evidence about didactic phenomena about which there are solid findings. Our sources of data are the results obtained from large-scale assessments. The debate on how a psychometric approach may help in defining the robustness of a solid finding in mathematics education is open and lively (Primi, 2017; for general discussions on the use of large-scale assessment data in research in mathematics education see Meinck, Neuschmidt & Taneva, 2017 and De Lange, 2007). Our approach is different from the ones outlined there and it works as follows: we start with a topic on which there are acknowledged solid findings, and we examine released items and results from large-scale libraries and datasets, searching for situations where there are pieces of evidence of these findings. Of course, the quantifications that we find are strictly dependent on the specific context where the data come from (as for every research result). Nevertheless, such collections of data in their whole, when interpreted with suitable theoretical lenses and integrated with information on the context, may give an articulated representation of the magnitude of the phenomenon observed, and possibly of local and related-to-context diversifications of it. Moreover, they can be used in mixed-method researches intended to deepen the understanding of the phenomena.

It must be underlined that having information on the extent of a problem is neither an explanation of the origin of the problem nor a solution to it. Nevertheless, we think that this kind of results may be useful both for teachers and for researchers. With a maybe naïf analogy, we might say that having epidemiological information on the diffusion of a disease is not directly related to the diagnosis or the treatment. Nevertheless, they are fundamental for driving the action of on-the-field physicians and for suggesting issues to medical researchers. In the same way, this kind of data may increase the awareness of teachers and help them in using the results of theoretical and empirical research in mathematics education. As it has been pointed out by Baccaglini-Frank and Di Martino (2017), a critical approach to large-scale assessment results may have a deep impact on teachers' practices and their professional development. "Critical" means also, in our view, "theoretically-driven". Moreover, we will show how new research issues may arise from them.

The case

The case that we present in this paper as the first example of this project is related to the loss of meaning in the use of algebraic symbols. The construction of algebraic language, and its relationship with the achievement of symbolic and formal thinking, is one of the fundamental shared keystones of mathematics curricula in all education systems worldwide- this has been

called *the dominance of the curriculum by algebra* (Greer, 2008). In Italy (the context of our case), in particular, this takes up much of the teaching time for mathematics teachers in Grade 9 and 10. The various steps towards mastering this construction, and the relative difficulties encountered, constitute a classic topic in mathematics education research and have been studied with the use of different theoretical and methodological lenses, mainly in qualitative terms, getting to a shared point which may be considered a solid finding: the use of symbols in algebra risks to become a *manipulative* activity on signs (letters and operational symbols), with no regard to what the letter and the symbols may stand for. The load of symbols is increasing along with the maths curriculum with heavy consequences (Bardini & Pierce, 2015), with deep discontinuities between school segments- typically, between primary and secondary schools, and between secondary and tertiary. In our context, this discontinuity happens mainly between Grade 8 and 9. This topic may also be considered a classical research topic in relationship with teachers' Mathematical Content Knowledge (Begle, 1972; Eisenberg, 1977; Strand & Mills, 2014).

Hence, this *loss of meaning* of the symbols used in algebra is a good situation for presenting the first example of our approach.

Framework and design of the study

A solid finding in mathematics education: loss of meaning in the use of algebraic symbols

As in many other countries worldwide, the first two years of second level-secondary schools in Italy (Grade 9 and Grade10) in all existing curricula include an incontestable key topic: algebra. The practice of algebra is one of the areas that demand the most time and attention of both teachers and students.

From a mathematical point of view, this topic poses many problems. The mathematical "objects" which are included, such as "literal expressions", "monomials" and "polynomials" (whatever the umbrella term "mathematical object" may include) are difficult, if not impossible, to define in a mathematical "rigorous" way at this scholastic level – for a more detailed analysis of this problem see Bolondi, Ferretti, and Maffia (2020). Pupils' learning, and above all their understanding of the meaning, refers to classroom practices regarding these objects, rather than to theoretical "definitions". Letters are used as symbols, and letters are designed as "indeterminate", "unknowns", parameters", "variables", and this refers to practices and contexts of use but the overall activity of students is the manipulation of symbols (see f.i. for researches focused on the context of this study Arzarello, Bazzini, & Chiappini, 1995; Chiarugi et al., 1995, or Bolondi et al., 2020 for a presentation of Italian textbooks approach to literal calculation). It must also be considered that the Italian cultural tradition based on the idealist philosophy (Bolondi, 2016) translates into classroom practices such that Italian high school students are asked mainly to learn theory and perform heavy algebraic calculations with the application of "rules", basically treatment transformation in the sense of Duval (1995): just to give an example, one of the most popular textbooks in Italy offers, in the literal calculation module alone, over 2000 exercises. Problem-solving and the use of mathematics to model real situations, even if recommended in National Guidelines, seems to be considered in many textbooks a secondary activity.

What is more, the international literature in mathematics education has highlighted in general how one of the main problems is the relationship that students create between symbols and their meaning (Kieran, 1992; Sfard, 1995; Bardini, Radford, & Sabena, 2005; Bardini, 2004) and the inadequacy of this relationship can often be linked to educational practices that insist on meaningless manipulations (Arzarello et al., 1995). Nevertheless, the use of letters for expressing relationships and for modelling begins very early in the school curriculum and is a classical research topic (see f.i. Doerr & Tripp, 1999; Dunne, 1998). As it has been pointed out, scholars generally agree that many students do not master the meaning of those symbols that they have learned to manage formally: classical references are Arcavi (1994, 2005), MacGregor and Stacey (1993) and Capraro and Joffrion (2006), and this has also been stated as a "solid" finding in Handbooks and Encyclopedias (Kieran, 1992, 2007, 2014).

The learning of algebra requires different cognitive components and interfaces the mathematical content, the use of language and semiotic representations, different forms of reasoning, meaning-making processes, and interpretations. Hence the loss of meaning of symbols in algebra can be interpreted and analysed with well-established theoretical lenses, such as Sfard's reification theory, Radford's objectification theory, and Duval's structural/functional approach. A theoretical approach coordinating these theories for interpreting the duality objects/processes is discussed in Santi, (2011a, 2011b).

A key step is the introduction of symbols and letters in the transition from arithmetic to algebra and in the use of "general" models of situations, but "it is little value using algebraic symbols in an answer when the person who posed the question may have little or no knowledge of their meaning" (Phillips, 2015). Arzarello et al. (2001), Linchevsky and Sfard (1992) and Sfard and Linchevsky (1994) highlight a continuity between the domain of arithmetic and the domain of algebra, where arithmetic provides operations and meanings that sustain algebraic thinking. Within the Reification Theory, arithmetical processes are reified in algebraic objects. There is a dialectic between the two domains that cannot be disregarded. On the one hand, without a leap into the algebraic mindset, algebra is a disguised form of arithmetic, which does not fully justify the introduction of symbolic language. On the other hand, without the link to arithmetical processes, algebra is perceived as meaningless by students and confined to manipulation of symbols that serve as empty signifiers.

The point is that algebraic symbols do not speak for themselves, therefore the algebraic competence may be identified with the ability to interpret the situation and act accordingly with versatility and adaptability in blending processes and objects appropriately. The difficulty with reification does not consist merely in the passage from process to object but from process to the process-object duality. If the signs are meaningless and the treatments of the expressions are performed mechanically, and the operational rules are meaningless operation, the students are unable to activate the necessary *interpretations*.

Mathematical thinking, mathematical concepts, meanings, and semiotics are distinguishable but inseparable within the mediated reflexive activity that intertwines them. Understanding the purpose of the system of algebraic signs (and more generally the symbolic language of maths), is one of the most difficult didactic problems of mathematics teaching/learning, linked in the

specific case of algebra to the role of model generalisation in the training of symbolic-mathematic thought processes (Radford, 2003).

As a synthesis, we may conclude as follows: if algebra is a language, then it should be analysed in its semantic as well as syntactic properties. The use of letters, and more generally of literal expressions that can then be manipulated, arose mainly from the need to generalise (for example, arithmetical properties of numbers) and that of modelling. In this sense, specifically for the age group considered, algebra plays an important role in the Framework of the OECD-Pisa survey (OECD, 2013). The semantic aspects are constructed early in the curriculum- for instance, when learning geometric formulas such as A=bxh. The role of letters in modelling processes is a crucial element of the curriculum before the learning of the rules for the treatment of expressions containing letters. There is a "meaning" of letters that is overshadowed by the overwhelming practices of manipulation of symbols. More generally, loss of meaning may occur when objects manipulated by the learners (in mathematics but also in Physics and other disciplines) are decontextualized (Brousseau, 1992; Johnston, Crawford & Fletcher, 1998). The loss of meaning of the letters for students all over the world- and in particular in Italy, the context of this study- can be considered a solid finding in mathematics education. There are powerful and well-established theoretical frameworks that can be used for describing the genesis and the outputs of this phenomenon, studies that make networking these theories, pieces of evidence from case studies all over the world, qualitative researches highlighting situations that can be considered exemplar models. What we want to add to this frame is an example of *quantification* at a systemic level of the phenomenon. This theoretically-driven analysis of quantitative data will also bring out particular features of the phenomenon itself that were not evident before. For instance, we will point out the spreading of the phenomenon all along the ability scale (we refer here to the latent trait measured by the large-scale assessment), thus showing how this approach may suggest new research issues.

Data for quantification: the context of INVALSI assessment tests and macro-level issues of learning

In this paper, we study the answers given by two cohorts of Italian students to two items in large-scale national assessment tests (the "INVALSI tests"). We have selected these items since, in an *a-priori* analysis of the library of INVALSI items, they seemed suitable for revealing and quantifying the loss of attribution of meaning to algebraic language, in the transition from middle school to high school. We remark that this is a cross-sectional study, i.e. two different cohorts of students at different points in their studies are assessed in the same year. Since our data comes from large-scale assessments of the whole population, we assume that the year-to-year variations are negligible.

The INVALSI tests are designed to gather and give back to pupils, teachers, and schools information and data about learning of mathematics at different steps- namely at Grades 2, 5, 8, 10, and 13. They highlight macro phenomena, sometimes extremely significant from a quantitative and statistical point of view. The released results moreover allow (particularly in the closed-ended question sections) the analysis of wrong choices, in order to better understand students' difficulties and their degree of diffusion in specific sub-categories of students (Bolondi, Ferretti & Giberti, 2018; Ferretti, Giberti & Lemmo, 2018). The items considered in this paper are taken from the tests administered in 2011 to students of Grades 8 and 10. The data considered

consist of the answers given by random samples of students of Grades 8 and 10 from across the country, involving 28,361 students in 1,312 classes representative of 586,790 students, and 50,838 students in 2,302 classes representative of 527,318 students respectively. The sample chosen for Grade 10 is larger since the architecture of the Italian school system is based on a unique path for all students till Grade 8, whilst from Grade 9 on there is a differentiation. The samples are representative with respect to geographical distribution, socioeconomic context, gender, typology of school (for Grade 10), and other variables (for a complete technical description of the INVALSI, 2011). The statistical analysis was carried out using the Rasch model (Rasch, 1960). This model belongs to the Item Response Theory (IRT) family. It is based on the assumption that the answer of a student to an item depends on his/her relative ability, i.e. the student's intrinsic ability as compared with the intrinsic difficulty of the item. The model scales both subjects and items along with the same *latent trait*, which in our case can be identified with mathematics ability, as defined in the mathematics framework which is the background of the INVALSI assessment.

All data, graphs, and tables provided here are taken from the freely accessible database Gestinv (www.gestinv.it). The complete original datasets of students' answers can be retrieved from the statistical servers of the INVALSI. Gestinv is a research tool widely used in Italian mathematics education researches (i.e. Ferretti, Gambini, & Santi, 2020; Ferretti & Bolondi, 2019).

The Items of the Spring: *a-priori* analysis

The case study that we present is a comparison of tasks where the question intent is focused on the meaning of an expression using letters, in a modelling situation. We point out that the physical situation is related to Hooke's law, but Italian students at these stages of their curriculum, generally, do not know an explicit formulation of Hooke's law. Hence the items can be considered as pure modelling tasks, where algebraic symbols are used in order to describe a situation expressed in words.

Using symbols (the letters L, L_0, K, P) for expressing physical quantities and symbols (the signs =, +, x) for expressing the relations between these quantities, is an example of using algebraic equality with reference to its "meaning". For instance, "multiplication by a constant" means "proportionality of the dilatation". This allows manipulating algebraically the expression in order to describe or derive physical properties, or to push further the algebraic symbolism, for instance by means of the tensor language. Entering the realm of algebra opens powerful perspectives, but this depends on maintaining a link to the meaning.

The first item was administered to Grade 8 students.

> Item 17
> The formula $L = L_0 + K\text{x}P$ expresses the length L of a spring according to the weight P applied. L_0 represents the length in centimetres of the "relaxed" spring; K represents how much the spring extends in centimetres when a unit of weight is applied.

Which of the following formulae best suits the following description:
"It is a very short and very hard spring (i.e. highly resistant to traction)"?

A. $L = 10 + 0.5 \times P$
B. $L = 10 + 7 \times P$
C. $L = 80 + 0.5 \times P$
D. $L = 80 + 7 \times P$

D17. La formula $L = L_0 + K \times P$ esprime la lunghezza L di una molla al variare del peso P applicato. L_0 rappresenta la lunghezza in centimetri "a riposo" della molla; K indica di quanto si allunga in centimetri la molla quando le si applica una unità di peso.
Quale delle formule elencate si adatta meglio alla seguente descrizione:
"*È una molla molto corta e molto dura (cioè molto resistente alla trazione)*"?

☐ A. $L = 10 + 0{,}5 \times P$

☐ B. $L = 10 + 7 \times P$

☐ C. $L = 80 + 0{,}5 \times P$

☐ D. $L = 80 + 7 \times P$

Figure 1: Item D17, INVALSI mathematics test for Grade 8, 2011, original formulation

The item is coherent with the Goals for skills acquirement and Learning Objectives as set out in the Italian National Guidelines (MIUR, 2012). The stimulus presents the general formula to express the length of a spring according to the weight applied and specifies the meaning of the symbols used. This is a simple literal expression (to use the common Italian textbook terminology) modelling mathematically the situation. To answer correctly, it is necessary to precisely interpret a statement in natural language and to understand the meaning of the symbols in the literal expression that translates the statement into mathematical language. The student must therefore carry out a typical conversion transformation (D'Amore 2006; Duval, 2008). This, then, is an item that in the conceptual framework of OCSE-Pisa is questioning on the linking processes between real and mathematical problems (Formulating and Interpreting; OECD 2013).

The second item was administered to Grade 10 students, and it is very similar to the previous one.

Item 24
The formula $l = l_0 + k \times P$ expresses the length l of a spring according to the weight P applied. l_0 represents the length in centimetres of the "relaxed" spring; k represents how much the spring extends in centimetres when a unit of weight is applied.
Which of the following formulae best suits the following description:

"It is a very long and very hard spring (i.e. highly resistant to traction)"?

A. $l = 15 + 0.5 \times P$
B. $l = 75 + 7 \times P$
C. $l = 70 + 0.01 \times P$
D. $l = 60 + 6 \times P$

D24. La formula $l = l_0 + k \cdot P$ esprime la lunghezza l di una molla al variare del peso P applicato. l_0 rappresenta la lunghezza in centimetri "a riposo" della molla; k indica di quanto si allunga in centimetri la molla quando si applica una unità di peso.
Quale delle formule elencate si adatta meglio alla seguente descrizione:
***"È una molla molto lunga e molto resistente alla trazione"*?**

☐ A. $l = 15 + 0.5 \cdot P$

☐ B. $l = 75 + 7 \cdot P$

☐ C. $l = 70 + 0.01 \cdot P$

☐ D. $l = 60 + 6 \cdot P$

Figure 2: Item D24, INVALSI mathematics test for Grade 10, 2011, original formulation.

The numeric values of the options change as compared with the Grade 8 item but the basic substance does not change, nor do the question intent or the meaning of the different options.

Results

Figure 3 shows the percentage of students' choices, while Figure 4 shows the main statistical data.

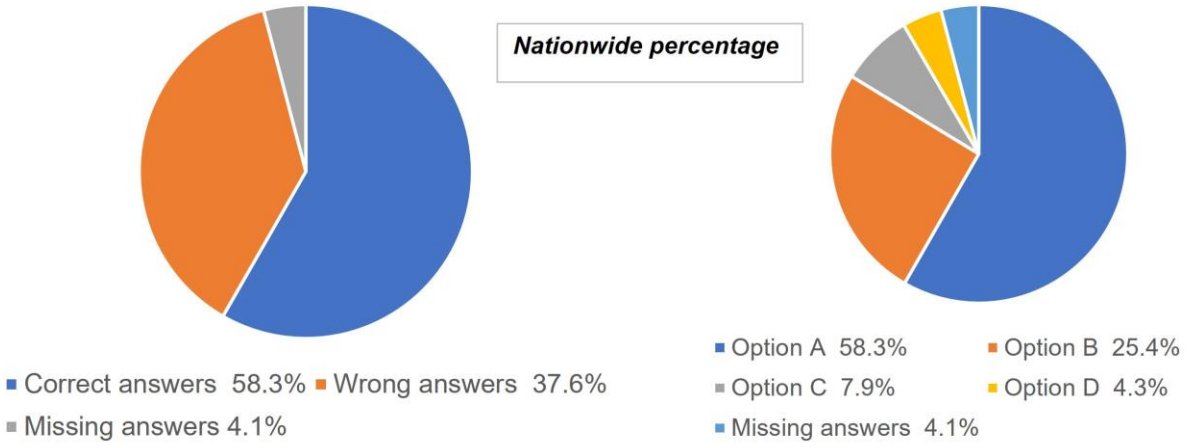

Figure 3: Percentage of choices, Item D17, INVALSI mathematics test of Grade 08, 2011

```
item:23 (D17)
Cases for this item  25892   Discrimination  0.43
Item Threshold(s):    -0.41   Weighted MNSQ   1.00
Item Delta(s):     -0.42
-------------------------------------------------------------------------------
 Label    Score     Count   % of tot  Pt Bis     t  (p)   PV1Avg:1 PV1 SD:1
-------------------------------------------------------------------------------
   1       1.00    15177      58.62    0.43    77.63(.000)  0.31     0.88
   2       0.00     6640      25.64   -0.25   -40.97(.000) -0.35     0.83
   3       0.00     2025       7.82   -0.17   -27.49(.000) -0.49     0.84
   4       0.00     1114       4.30   -0.17   -28.50(.000) -0.70     0.84
   7       0.00      217       0.84   -0.06    -9.59(.000) -0.57     0.82
   9       0.00      719       2.78   -0.12   -19.98(.000) -0.63     0.84
===============================================================================
```

Figure 4: Statistical data, Item D17, INVALSI mathematics test of Grade 8, 2011

Overall, the psychometric performance of this item was very good, in particular the fitting with the model. The Item delta level is -0.42 (on a scale centered in 0), so the item can be qualified of medium-easy difficulty. As we can see from the graphics in Figure 3, the correct response was supplied by 58.3% of students. One of the most interesting facts that emerge from the data is that 25.4% of the total (almost 70% of students who made a mistake) chose option B. Thus, they correctly interpreted the fact that the spring is very short but made a mistake in evaluating the role and the meaning of coefficient K.

The distractor plots are very useful tools for examining more deeply these outputs. In the plot in Figure 5, the x-axis shows the students' level of ability, measured overall by the test. The continuous line is the curve of answers as elicited by the model, while the dotted lines are the empirical data (percentages of answers collected).

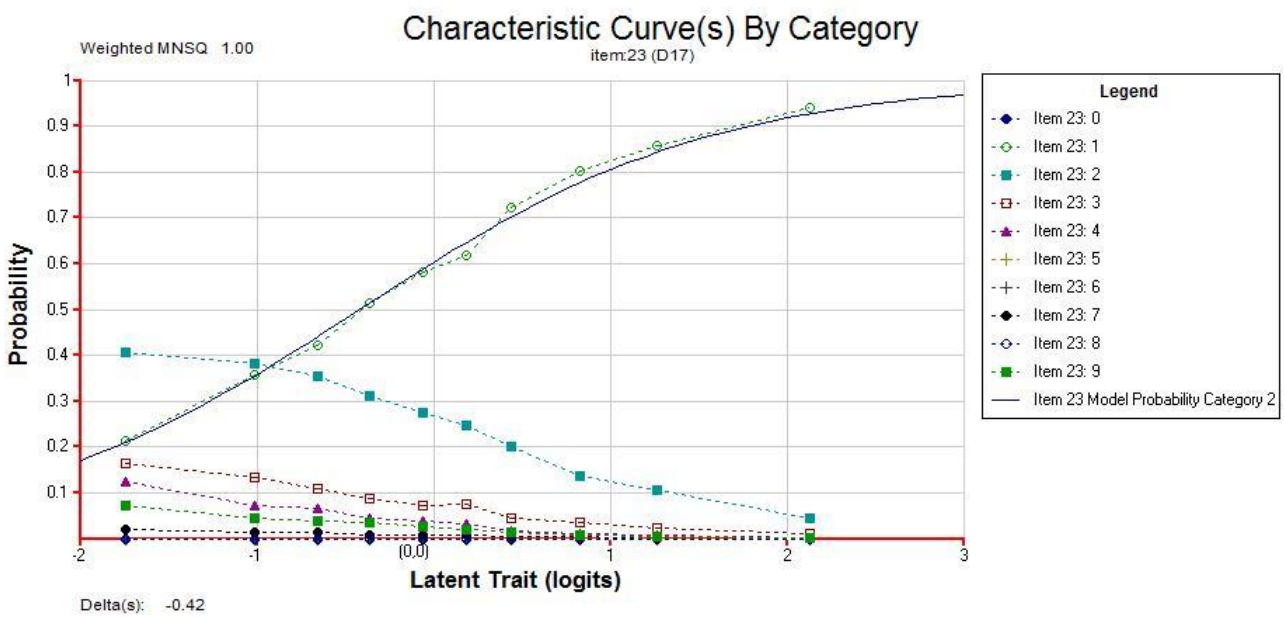

Figure 5: Distractor Plot of Item D17, INVALSI mathematics test of Grade 8, 2011

This distractor plot shows how the students who identified the first parameter and not the second (option B answer, labelled as Item 23.2 line) are mainly low-performance students; the frequency of this choice rapidly decreases as the ability level rises.

The results of the item of Grade 10 are shown below; Figures 6 and 7 show the percentage of answers and some statistical data.

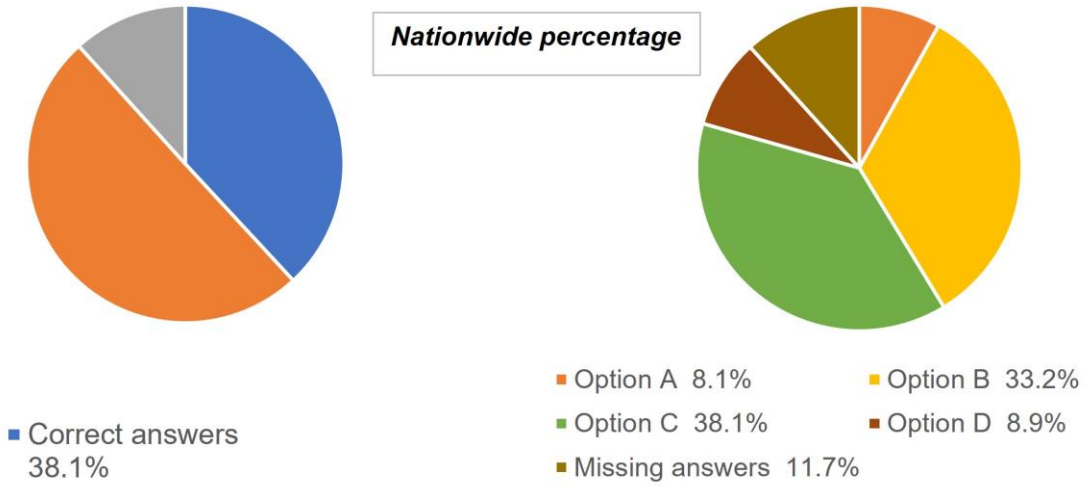

Figure 6: Percentage of answers, Item D24, INVALSI mathematics test for Grade 10, 2011

```
item:42 (MD24)
Cases for this item  43458   Discrimination  0.39
Item Threshold(s):     0.53   Weighted MNSQ   1.05
Item Delta(s):      0.53
-------------------------------------------------------------------------------
 Label    Score     Count   % of tot  Pt Bis     t  (p)   PV1Avg:1 PV1 SD:1
-------------------------------------------------------------------------------
   1       0.00     3481       8.01   -0.12   -24.83(.000) -0.37     0.79
   2       0.00    14432      33.21   -0.13   -27.01(.000) -0.17     0.86
   3       1.00    16892      38.87    0.39    87.03(.000)  0.45     1.02
   4       0.00     3944       9.08   -0.11   -23.38(.000) -0.32     0.81
   7       0.00      117       0.27   -0.04    -8.61(.000) -0.78     1.13
   9       0.00     4592      10.57   -0.20   -42.22(.000) -0.56     0.87
===============================================================================
```

Figure 7: Statistical data, Item D24, INVALSI mathematics test for Grade 10, 2011

In this case, too, the measurement parameters are valid: the Item delta level, however, is 0.53 so the item positions itself in the medium-difficult range, within the context of the test. Only 38.1% of the students supplied the correct answer, while the number of missing answers is 7% higher than that of the previous question. We remark that 11,7% of missing and invalid answers is a very high percentage for a closed-ended item. The most chosen option was option B, where the

coefficient of dilatation is obviously wrong. Analysis of the Distractor Plot (Figure 8) also reveals that students of medium and medium-high academic level opted for this option.

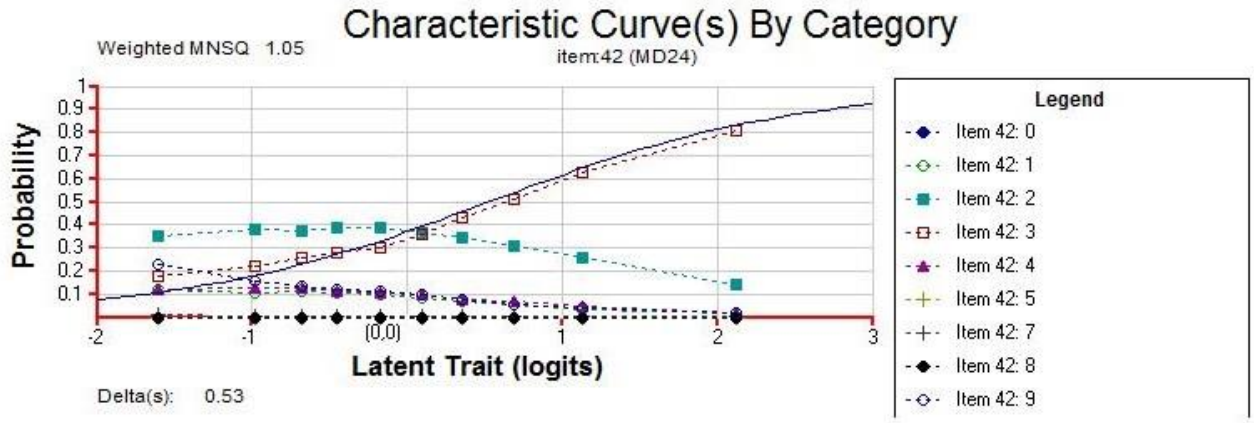

Figure 8: Distractor plot, Item D24, INVALSI mathematics test for Grade 10, 2011

Discussion

The analysis of the two items administered in Grade 8 and Grade 10 respectively, shows that semantic control of literal expressions (used as a tool of generalisation and modelling) decreases significantly between the two educational levels: the number of correct answers supplied decreased from 58.3% to 38.1%. This is not, of course, a longitudinal or even pseudo-longitudinal result (Baltagi, 2011); nevertheless, it describes the behaviour of representative samples of two cohorts of students in the same national context. From the statistical point of view, this increase of difficulty is visible through the value of the Item delta inside the tests, which jumps from -0.42 to 0.53. Both items have good psychometric features, they discriminate between students and fit with the model. The analysis of the items shows that the question intent is the same, and the verbal formulation is substantially the same. The different options have similar features.

In the first item, two options (the correct one, A, and option B) describe correctly at least the length of the spring, and these options collect globally 83.7% of the students' answers. Two grades later, three options describe correctly the length of the spring; nevertheless, even with three possibilities, only 80.2% of students choose them.

Despite two school years more, devoted mainly (in Italian schools) to the development of algebraic manipulation, there is a surprising difference in the percentage of correct answers: Grade 10 students provided 20% fewer correct answers than Grade 8's. "Missing" students are students of all levels of ability, even high. The distractor plots allow the drafting of an initial analysis regarding which students knew how to respond correctly to the item and to identify sub-groups of students who chose a distractor answer with significant frequency. For instance, in the eighth decile, the percentage of correct answers slumps from 80% to 50%; the percentage of correct answers in the ninth decile decreases from 85% to approximately 60%. Symmetrically, in Grade 8 starting from the fourth decile, more than 50% of the population gives the correct

answer; in Grade 10 this happens starting from the eighth decile. The most chosen wrong option is the preferred one, in Grade 10, for the first 5 deciles, hence for half of the population. Missing answers (which usually are very few in a multiple-choice test with no penalties for wrong answers, as in this case) jump from 4% to 11%: this too can be interpreted as difficulty in giving a meaning to the elements of the task.

Conclusion and further issues

We started from the "loss of meaning" in algebra as an exemplary case of "solid finding" in mathematics education, in the sense of EMS (2011). The theoretically-driven use of large-scale assessment data-sets allowed quantification of this effect of loss of meaning of algebraic symbols (which is a well-documented and explored phenomenon in the literature) in the Italian context, and to identify which categories of students this phenomenon affects in particular. Our approach moreover highlighted the fact that this phenomenon impacts also medium- and high-level students: as we have seen, there is a dramatic decrease of the percentage of correct answers for students graded in the top deciles (with respect to the overall ability measured by the test). Previous qualitative researches and epistemological discussions showed that this phenomenon can be related to deeply rooted didactic habits which put in the centre of students' activity the manipulation of symbols, disregarding what a symbol stands for. The combination of qualitative results, quantitative evidence, and theoretical discussions suggests further issues, both for researchers and for teachers to adopt in their teaching practices. In general, following Brousseau (1992), our case confirms with quantitative evidence that loss of meaning may occur when learning is decontextualised and, for algebra, syntactic treatments take over semantic attentions. Restoring a sense to algebra (apart from its role as a didactic tool) confirms to represent a key challenge, both for researchers and teachers. Teachers should be aware that this phenomenon is quantitatively significant. The same challenge may occur in other disciplines, especially when the "mathematization" of the discipline is relevant.

This paper is a specific case study to exemplify how the results of large-scale assessment projects can contribute, possibly within a mixed-method research approach (Hart et al., 2009, Johnson & Onwuegbuzie, 2004) and integrated with solid findings of education research (Kieran, 2014), to a better understanding of a didactic phenomenon. The possibility of analysing huge quantities of data, gathered through large-scale assessment activities and stored in digital databases, allows a more detailed and quantitative description (and possibly new directions of investigation) of results obtained through classical research methods, and to verify previous theoretical descriptions.

Acknowledgements
This work was supported by the Open Access Publishing Fund of the Free University of Bozen-Bolzano.

Ethical statement
This research has been performed in accordance with the Code of Ethics of the Free University of Bozen (https://www.unibz.it/assets/Documents/University/unibz-code-of-ethics-it.pdf) and controlled by the Ethics Committee of the University.

References

Arcavi, A. (1994). Symbol Sense: Informal Sense-making in Formal Mathematics. *For the learning of Mathematics, 14*(3), 24-35.

Arcavi, A. (2005). Developing and using Symbol Sense in Mathematics. *For the Learning of Mathematics, 25*(2), 42-47.

Arzarello, F., Bazzini, L., & Chiappini, G. (1995). The construction of algebraic knowledge: towards a socio-cultural theory and practice. In L. Meira & D. Carraher (Eds), *Proc. 19th Conf. of the Int. Group for the Psychology of Mathematics Education*, (Vol.1, pp.119-134), Recife: PME.

Arzarello, F., Bazzini, L., & Chiappini, G. (2001). A model for analysing algebraic processes of thinking. In R. Sutherland, T. Rojano, A. Bell, & R. Lins (Eds.), *Perspectives on school algebra* (pp. 61-82). Dordrecht, The Nederlands: Kluwer.

Baccaglini-Frank, A., & Di Martino, P. (2017). Beyond performance results: analyzing the informational and developmental potentials of standardized mathematics tests. *For the Learning of Mathematics, 37(3)*, 39-44.

Baltagi, B. (2011). Panel Data Analysis. In P. Lavrakas (Ed.), *Encyclopedia of Survey Research Methods* (pp. 568-569). Thousand Oaks: Sage Publications, Inc.

Bardini, C. (2004). Le rapport au symbolisme algébrique: une approche didactique et épistémologique. In C. Castela & C. Houdement (Eds.), *Actes du séminaire national de didactique des mathématiques* (pp. 141-166). Paris: Ed. ARDM et IREM de Paris7.

Bardini, C., Radford, L., & Sabena, C. (2005). Struggling with variables, parameters and indeterminate objects or how to go insane in mathematics. In H. L. Chick, & J. L. Vincent (Eds.), *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education,* (Vol.2, pp. 129-136). Melbourne: PME.

Bardini, C., & Pierce, R. (2015). Assumed Mathematics Knowledge: the Challenge of Symbols. *International Journal of Innovation in Science and Mathematics Education, 23(1),* 1-9.

Begle, E. G. (1972). Teacher Knowledge and Student Achievement in Algebra. *SMSG Reports, No. 9*. Stanford: School Mathematics Study Group.

Bolondi, G. (2016). Epistemology and didactics in Federigo Enriques. *Isonomia-Epistemologica, Teaching and learning Mathematics, vol.8,* 19-30.

Bolondi, G., Ferretti, F., & Giberti, C. (2018). Didactic Contract as a Key to Interpreting Gender Differences in Maths. *Journal of Educational, Cultural and Psychological Studies (ECPS Journal), 18*, 415-435.

Bolondi, G., Ferretti, F., & Maffia, A. (2020). Monomials and polynomials: the long march towards a definition. *Teaching Mathematics and its Applications, 39*(1), 1-12.,

Bosch, M., Dreyfus, T., Primi, C., & Shiel, G. (2017). *Solid findings in mathematics education: What are they and what are they good for?* In T. Dooley & G. Gueudet, G. (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education (CERME10, February 1-5, 2017),* (Vol.1, pp.53-56). Dublin, Ireland: DCU Institute of Education and ERME.

Brousseau,G. (1992). "Didactique": What it can do for the teacher. *Research en Didactique des Mathematiques, 7,* 7-40.

Capraro, M. M., & Joffrion, H. (2006). Algebraic Equations: Can Middle-School Students Meaningfully translate from Words to Mathematical Symbols? *Reading Psychology, 27*, 147-164.

D'Amore, B. (2006). Oggetti matematici e senso. Le trasformazioni semiotiche cambiano il senso degli oggetti matematici [Semiotic transformations change the meaning of mathematical objects]. *La matematica e la sua didattica, 20*(4), 557-583.

De Lange, J. (2007). Large-Scale Assessment and Mathematics Education. In F. K. Lester, Jr. (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 1111–1142). Charlotte, NC: Information Age Publishing.

Doerr, H., & Tripp. J. (1999). Understanding how students develop mathematical models. *Mathematical Thinking and Learning*, 1(3), 231–254.

Dreyfus, T. (2017). What are solid findings in mathematics education? In T. Dooley & G. Gueudet, G. (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education (CERME10, February 1-5, 2017),* (Vol.1, pp.57-62). Dublin, Ireland: DCU Institute of Education and ERME.

Dunne, T.A, (1998). Mathematical modelling in years 8 to 12 of secondary schooling. In P. Galbraith, W. Blum, G. Booker and I.D. Huntley (Eds.), *Mathematical Modelling: Teaching and Assessment in a Technology-Rich World*, (pp. 29-37). Chichester: Horwood Publishing.

Duval, R. (1995). *Sémiosis and pensée humaine.* Bern: Peter Lang.

Duval, R. (2008). Eight Problems for a Semiotic Approach in Mathematics Education. In L. Radford, G. Schubring, & F. Seeger (Eds.), *Semiotics in Mathematics Education: Epistemology, History, Classroom, and Culture* (pp. 39-61). Rotterdam: Sense Publishers.

Eisenberg, T. H. (1977). Begle Revisited: Teacher Knowledge and Student Achievement in Algebra**.** *Journal for Research in Mathematics Education*, *8*(3), 216-222.

EMS (2011). "Solid Findings" in Mathematics Education. *Newsletter of the European Mathematical Society*, *81*, 46-49.

EMS (2012). What are the Reciprocal Expectations between Teacher and Students? Solid Findings in Mathematics Education on Didactical Contract. *Newsletter of the European Mathematical Society*, *82*, 53-55.

EMS (2013). Sociomathematical Norms: In Search of the Normative Aspects of Mathematical Discussions. *Newsletter of the European Mathematical Society*, *88*, 59-61.

EMS (2015). Solid Findings: Students' Over-reliance on Linearity. *Newsletter of the European Mathematical Society*, *95*, 51-53.

Ferretti, F., Bolondi, G. (2019). This cannot be the result! The didactic phenomenon "the Age of the Earth". *International Journal of Mathematical Education in Science and Technology. DOI:*10.1080/0020739X.2019.1670366

Ferretti, F., Gambini, A., & Santi, G. (2020). The Gestinv Database: A Tool for Enhancing Teachers Professional Development within a Community of Inquiry. In H. Borko and D. Potari (Eds.), *Proceedings of the Twenty-fifth ICMI Study School Teachers of mathematics working and learning in collaborative groups* (pp.621–628). Portugal: University of Lisbon.

Ferretti, F., Giberti, C., & Lemmo, A. (2018). The Didactic Contract to Interpret some Statistical Evidence in Mathematics Standardized Assessment Tests. *EURASIA Journal of Mathematics, Science and Technology Education*, *14*(7), 2895-2906.

Greer, B. (2008). Algebra for all? *The Mathematics Enthusiast*, *5*(2), 423-428.

Hart, L.C., Smith, S.Z., Swars, S. L., & Smith, M.E. (2009). An Examination of Research Methods in Mathematics Education (1995-2005). *Journal of Mixed Methods Research*, *3*(1), 26-41.

INVALSI (2011). *Servizio Nazionale di Valutazione 2010-11. Le rilevazioni degli apprendimenti A.S. 2010/11*. Roma: Invalsi.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, *33*(7), 14-26.

Johnston, I. D., Crawford, K. & Fletcher, P. R. (1998). Student difficulties in learning quantum mechanics, *International Journal of Science Education, 20:4*, 427-446,

Kieran, C. (1992). The learning and teaching of school algebra. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp.390-419). New York: MacMillan.

Kieran, C. (2007). Learning and teaching algebra at the middle school through college levels. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning: a project of the National Council of Teachers of Mathematics*, (Vol. 2, pp. 707-762). Charlotte, NC: Information Age Pub.

Kieran, C. (2014). Algebra Teaching and Learning. In S. Lerman (Ed), *Encyclopedia of Mathematics Education*, (pp. 27-32). Dordrecht: Springer.

Linchevsky, L., & Sfard, A. (1992). Rules without reasons as processes without objects – the case of equations and inequalities. In Furinghetti, F. (Ed.), *Proceedings of the 15th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 317-324). Assisi, Italy: PME.

MacGregor, M., & Stacey, K. (1993). Cognitive Models Underlying Students' Formulation of Simple Linear Equations. *Journal for Research in Mathematics Education*, *24*, 217–232.

Meinck, S., Neuschmidt, O., & Taneva, M. (2017). Workshop Theme: "Use of Educational Large-Scale Assessment Data for Research on Mathematics Didactics". In G. Kaiser (Ed.), *Proceedings of the 13th International Congress on Mathematical Education*, (pp.741-742). ICME-13 Monographs. Cham: Springer.

MIUR (2012). Indicazioni Nazionali per il curricolo della scuola dell'infanzia e del primo ciclo di istruzione [National Guidelines for the curriculum in Kindergardens and Primary schools]. *Annali della pubblica istruzione*, vol. 88, numero speciale 2012. Firenze: Le Monnier.

Niss, M. (2012). Models and Modelling in Mathematics Education. *Newsletter of the European Mathematical Society*, *86*, 49-52.

OECD. (2013). *Assessment and Analytical Framework: Mathematics, Reading, Science*. OECD Publishing. Retrieved December 12, 2019, from http://dx.doi.org/10.1787/9789264190511-en.

Phillips, C.G. (2015). An Improved Representation of Mathematical Modelling for Teaching, Learning and Research. *International Journal of Innovation in Science and Mathematics Education*, *23*(4), 51-63.

Primi, C. (2017). *Solid findings in mathematics education: A psychometric approach.* In T. Dooley & G. Gueudet, G. (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education (CERME10, February 1-5, 2017),* (Vol.1, pp.63-67). Dublin, Ireland: DCU Institute of Education and ERME.

Radford, L. (2003). Gestures, Speech, and the Sprouting of Signs: A Semiotic-Cultural Approach to Students' Types of Generalization. *Mathematical Thinking and Learning, 5*(1), 37-70.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.

Santi, G. (2011a). Objectification and semiotic function. *Educational Studies in Mathematics, 77*, 285-311.

Santi, G. (2011b). Meaning of mathematical objects: a comparison between semiotic perspectives. In M. Pytlak, T. Rowland, & E. Swoboda (Eds), *Proceedings of the Seventh Congress of the European Society for Research in Mathematics Education (CERME-7)* (pp. 2503-2512). Poland: University of Rzeszów.

Sfard, A. (1995). The Development of Algebra: Confronting Historical and Psychological Perspectives. *The Journal of Mathematical Behavior, 14*(1), 15-39.

Sfard, A., & Linchevski, L. (1994). The Gains and the Pitfalls of Reification: The Case of Algebra. *Educational Studies in Mathematics, 26,* (2-3), 191-228.

Strand, K., & Mills, B. (2014). Mathematical Content Knowledge for Teaching Elementary Mathematics: A Focus on Algebra. *The Mathematics Enthusiast, 11*(2), 385-432.