



HAL
open science

How do primary teachers interpret and use standardized assessment: the case of the crochet placemats

Eleonora Faggiano, Federica Ferretti, Ferdinando Arzarello

► To cite this version:

Eleonora Faggiano, Federica Ferretti, Ferdinando Arzarello. How do primary teachers interpret and use standardized assessment: the case of the crochet placemats. Twelfth Congress of the European Society for Research in Mathematics Education (CERME12), Feb 2022, Bozen-Bolzano, Italy. hal-03753407

HAL Id: hal-03753407

<https://hal.science/hal-03753407>

Submitted on 18 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

How do primary teachers interpret and use standardized assessment: the case of the crochet placemats

Eleonora Faggiano¹, Federica Ferretti² and [Ferdinando Arzarello](mailto:ferdinando.arzarello@unito.it)³

¹University of Bari Aldo Moro, Italy; eleonora.faggiano@uniba.it

²University of Ferrara, Italy; federica.ferretti@unife.it

³University of Torino, Italy; ferdinando.arzarello@unito.it

This paper presents the first findings from a survey, administered to 421 Italian in-service primary teachers, on their beliefs regarding the knowledge and skills investigated by the national standardized assessment (INVALSI) tests, their proximity to didactic practices in Mathematics and the role they assume within the school context. The case presented in this paper is discussed in order to investigate the way teachers interpret data coming from standardized assessment and if/how they use them in their teaching practice. Findings show an overspread meta-didactic conflict generated by teachers' difficulties in interpreting INVALSI tests and in using them coherently with the framework on which the tests have been designed.

Keywords: Standardized assessment, meta-didactic conflict, primary teachers.

Rationale and theoretical framework

At international level, the strictly link between standardized assessment and mathematics education is increasingly emerging (De Lange, 2007). Standardized assessment has the main purpose of system evaluation and its repercussions within social, political and educational fields, is increasingly taking hold, both at the level of the school system, and at the level of impact on classroom practices (Looney, 2011). There is a growing interest in broadening standardized assessment beyond the evaluation of school systems to mathematics education research as a new methodological tool (De Lange, 2007; Meinck, Neuschmidt, & Taneva, 2017). In Italy, the INVALSI national agency is responsible of the standardized mathematics assessment (INVALSI tests) that occurred, for the first and second cycle of education, since 2008. The INVALSI tests are administered at a census nationally and the results, elaborated on a valid statistical sample, are returned to the schools and publicly discussed every year. The theoretical framework of the INVALSI tests is in line with the main research results in mathematics education and with the National Curriculum Guidelines (INVALSI, 2018). This ensures that the macro-phenomena highlighted in the standardised assessment can be framed with the main research lenses and that they highlight new characterisations or phenomena (Bolondi, Ferretti, & Santi, 2019, Ferretti, & Bolondi, 2019, Ferretti, & Gambini, 2018). Several researches show how the results of the standardized assessment can be used in a formative perspective within didactic practices and as a tool within teachers' professional development paths (i.e., Doig, 2006; Di Martino, & Baccaglini-Frank, 2017; Ferretti, Gambini, & Santi, 2020). The research project within which this study moves, is in tune with these strengths of thought and aims to identify training needs of Italian teachers and to propose guidelines for the improvement of teaching practices, regarding the use of mathematics INVALSI tests. In particular, we investigate the presence of any conflicts, in the sense of Sfard (2008), that might exist between the "language of standardised assessments" and

mathematics teachers' interpretation of it (Arzarello, & Ferretti, 2021). As we will explain in the following section, one of the main aims of this research is to understand the role and meanings that teachers attribute to the INVALSI tests, investigating how they interpret, consider and use the INVALSI tests and their results.

The interdisciplinary research project and our research questions

Herein, we analyse the first results of an interdisciplinary research project aimed at investigating the link between the INVALSI tests in Mathematics with the teaching and learning processes of Mathematics, in particular with teaching practices. The research project is conducted by the INVALSI Disciplinary Didactics Group of the Italian Society of Research in Didactic S.I.R.D., composed by mathematics education researchers and pedagogists. The interdisciplinary collaboration consisted in the design and in the administration of a tool for detecting teachers' attitudes towards INVALSI agency with its aims and working methods, as well as its tests in Mathematics and their repercussions on teaching practices. We are interested to understand what are the “tools” possessed and used by the teachers: to read and interpret INVALSI tests and their results; to identify possible effects of the INVALSI tests on their Mathematics teaching practices. The analysis of data coming from the results of the INVALSI standardized assessment is not enough to identify training needs in this direction at a national level within schools and to propose guidelines for the improvement of practices regarding the use of INVALSI tests. Rather, to reach these goals, we intend to develop a meta level analysis focusing on the way teachers interpret and use these data in their teaching practices. To do that, a survey was designed and administered to investigate primary teachers' beliefs regarding the knowledge and skills investigated by the INVALSI tests, their proximity to didactic practices in Mathematics and the role they assume within the school context.

In this paper we analyse data concerning the Mathematics Education section of the survey. The section focuses on 7 tasks of the INVALSI tests, chosen because considered suitable to highlight different didactic macro phenomena. With respect to each of the 7 tasks, teachers are asked to answer questions aiming to investigate: their ability to read INVALSI data and identify the reasons of students' mistakes; how suitable an INVALSI task is considered for assessing students learning and how commonly it is used in assessment practices. The final part of the Mathematics Education section of the survey contains some transversal questions that aim at investigating how the mathematical contents and the skills detected with the INVALSI tests are more or less close to the daily personal didactic practices and perceived as consistent/inconsistent with the National Curricula Guidelines. All these variables represent the focus of the survey and are the object of the analysis proposed in this study. In what follows we will exemplify our study presenting and analysing teachers' responses to some of the questions of the survey. More specifically, we intend here to answer the following research questions:

Q1: to what extent teachers' interpretation of the INVALSI tests and their results is coherent with the framework on which the tests have been designed?

Q2: to what extent the mixed (qualitative and quantitative) investigation based on our survey results to be a tool to answer Q1?

The crochet placemats task

This paper focuses on some questions of our survey, which refer to the INVALSI task of Figure 1.

<p>To make 4 crochet placemats, grandmother uses 6 balls of cotton.</p> <p>a) How many balls of cotton of the same type she has to use to make 20 placemats?</p> <p>Answer:.....</p> <p>b) Write how you did to find the answer.</p>
--

Figure 1: Task 11, Mathematics INVALSI test Grade 05, s.y. 2012-13

Teachers were given a set of students' answers to the task, consisting in both the numerical result and the solution strategy. For each of the proposed student answers, teachers were asked to choose its degree of correctness (*Completely Incorrect*, *Partially Incorrect*, *Mainly Correct*, *Completely Correct*). These 4 possible options are given to the teachers without an explanation of the difference between *Partially Incorrect* and *Mainly Correct*: to the teacher's choice of *Partially Incorrect*, we attributed the feeling that what is important for the teachers in that answer of the student is the fact that it is incorrect (even if only partially), rather *Mainly Correct* is the answer that we expected to be chosen by the teachers which consider the answer of the student to be correct (even if not completely).

The proposed student answers are:

- 1) 30, I calculated $20+6+4$
- 2) 24, I multiplied 6 balls by 4 placemats, and I got 24
- 3) 120, I have multiplied 6 balls by 20 placemats
- 4) 30, Since for 4 placemats we need 6 balls of cotton (so 2 more), for 20 placemats we just need to do plus 10
- 5) 30, The grandmother uses a ball and a half to prepare a placemat, so I did this $= 20 \times 1.5 = 30$
- 6) 30, To do 20 placemats I have to do 5 times 4 placemats, so I need 30 balls
- 7) 24, I have multiplied 20 placemats by 1 ball and a half, and I got 24.

As we can notice, option 1) reports a correct numerical answer supported by a wrong strategy; on the contrary, option 7) reports a correct strategy with a wrong numerical result. Only the options 4), 5) and 6) are correct. Despite the solution strategies reported in these three options are all correct, they are characterized by different approaches. Options 5) and 6) mainly follow a multiplicative model of reasonings, whereas option 4) makes use of a mixed additive-multiplicative model (Arzarello, 2018).

Results

The survey, that has been offered to be voluntarily filled on-line, involved 421 Italian in service primary teachers: most of them teach since more than ten years.

Table 1 shows the valid percent of the teachers' responses inherent to the wrong student answers proposed in the survey. A first look at the responses given by the teachers shows that most of them

(more than 80%) recognised the first three options to be incorrect (Table 1). In particular, it seems to be reasonable that the correct numerical result (30) of the student answer, even in presence of a wrong strategy, can justify the fact that the number of teachers, who marked option 1) as *Completely Incorrect* is slightly lower with respect to those who considered *Completely Incorrect* options 2) and 3). Conversely, student answer 7) was considered *Incorrect (Completely or Partially)* by the 68.9% of the teachers (Table 1), even in presence of a correct strategy. A deeper analysis of data reveals that the 90% of the teachers which evaluated option 1) to be *Correct (Completely or Mainly)*, considered *Completely Incorrect* option 7), and at the same time, the 99% of those who evaluated option 7) to be *Correct (Completely or Mainly)*, considered *Incorrect (Completely or Partially)* option 1). This inverse correlation between teachers' responses given to student answers 1) and 7) is also confirmed by the factorial analysis we will present below (see Table 2).

Table 1: Valid percent of teachers' responses

Solution	Valid Percent			
	Completely incorrect	Partially incorrect	Mainly correct	Completely correct
1) 30, I calculated $20+6+4$	81.9	14.9	2.6	0.6
2) 24, I multiplied 6 balls by 4 placemats, and I got 24	84.9	11.9	2	1.2
3) 120, I have multiplied 6 balls by 20 placemats	85.5	12.5	1.2	0.8
4) 30, Since for 4 placemats we need 6 balls of yarn (so 2 more), for 20 placemats we just need to do plus 10	23	32	34.8	10.2
5) 30, The grandmother uses a ball and a half to prepare a placemat, so I did this = $20 \times 1.5 = 30$	4.7	5.2	18.8	71.3
6) 30, To do 20 placemats I have to do 5 times 4 placemats, so I need 30 balls	6.9	8.3	21.4	63.4
7) 24, I have multiplied 20 placemats by 1 ball and a half, and I got 24	46.3	22.6	29.7	1.4

Concerning the other three given answers, number 5) is mostly considered *Completely Correct*. A high percentage of teachers also recognised solution strategy 6) to be *Completely Correct*, while the teachers' responses inherent to solution strategy 4) reveal to be much noisier than numbers 5) and 6).

Also for the three correct student answers, some interesting results can be found observing data more in details. First of all, it seems worth of note that among the 192 teachers (45.6% of the total) which considered both the options 5) and 6) to be *Completely Correct*, only 29 also saw option 4) as *Completely Correct*. That is only the 6.9% of the teachers considered all the three student answers 4), 5 and 6) to be *Completely Correct*. Among the teachers who considered *Completely Correct* both options 5) and 6), we have also found out the number of those who saw option 4) as at least *Partially Incorrect*: they are 82 (the 19.5% of the total). This means that the 42.7% of those teachers, who considered *Completely Correct* answers 5) and 6), judged *Incorrect (Completely or Partially)* the mixed solution strategy presented in answer 4). Conversely, focusing on the total number of teachers

(232), who considered *Incorrect (Completely or Partially)* option 4), we can observe that 82 of them (that is the 35.3%) judged *Completely Correct* student answers 5) and 6).

The bivariate factor analysis (Suhr, 2006), using the Varimax method (SPSS software), allows the identification of two factors that saturate 44% of the total variance (Table 2). These results are interesting: in fact, option 1) and option 7) are in the same component with opposite signs; option 3) and option 2), both reporting wrong multiplicative model of reasoning, are in the same component; the first component also contains the two correct options 5) and 4) with opposite signs with respect to options 3) and 2); also the two most recognised correct strategies in options 5) and 6) belong to the the same component. However, some options reveal low factorial coefficients, so highlighting the presence of some critical aspects. These critical aspects are worthwhile of a further investigation, which we are now pursuing using qualitative methods of analysis.

Table 2: Varimax Rotated Factor Matrix with Kaiser normalization applied to teachers' responses

	Components	
3) 120, I have multiplied 6 balls by 20 placemats	.721	
2) 24, I multiplied 6 balls by 4 placemats, and I got 24	.715	
5) 30, The grandmother uses a ball and a half to prepare a placemat, so I did this = $20 \times 1.5 = 30$	-.567	.508
4) 30, Since for 4 placemats we need 6 balls of yarn (so 2 more), for 20 placemats we just need to do plus 10	-.424	
7) 24, I have multiplied 20 placemats by 1 ball and a half, and I got 24		.751
1) 30, I calculated $20+6+4$		-.490
6) 30, To do 20 placemats I have to do 5 times 4 placemats, so I need 30 balls		.449

Teachers were also asked to answer two other questions about their awareness of the validity of the task in order to assess students' learning and the extent of the use they make of similar tasks in their assessment test. Diagrams in Figure 2 show the valid percent of teachers' answers to these questions.

As those above, also these data show some unexpected correlation. We will discuss this issue in the next Section.

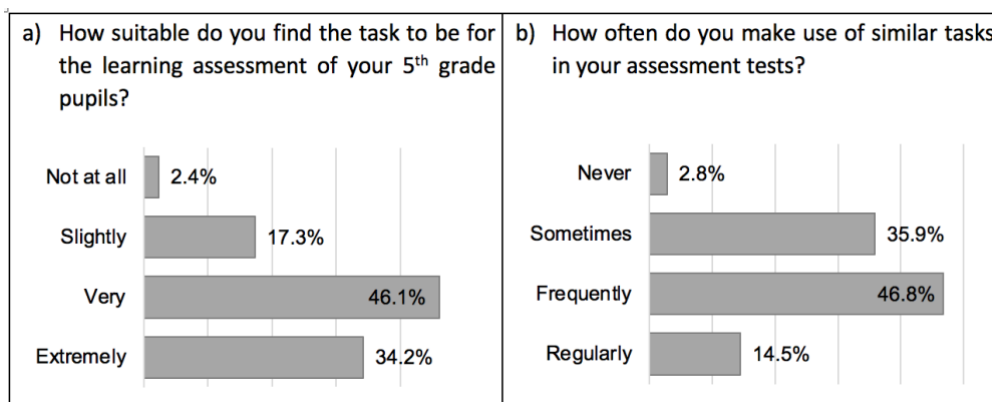


Figure 2: Teachers' responses to the questions about the validity of the task to assess students' learning and the use of similar questions in their assessment test

Discussion

We have already pointed out some contradictory aspects within teachers' answers to our survey. Figure 2 is emblematic for that. In fact, with respect to the two questions in it, we can see that, for the 80.3% of the teachers, the crochet placemats task is (*Very* or *Extremely*) suitable for the students' learning assessment, while the 61.3% of the teachers declared to make use (*Frequently* or *Regularly*) of similar questions in their assessment test. A first possible interpretation of this discrepancy might be connected with the perception of a higher difficulty embedded in the assessment of argumentative and problem solving skills, especially using open questions: although teachers recognize that the crochet placemats task is suitable to assess students' learning, in their usual assessment tests they prefer to assign more procedural exercises. However, we believe that the discrepancy highlighted by our data can be also considered in the light of the mentioned teachers' difficulty in recognising the correctness of “unusual” solution strategies.

We can interpret all these results basing on the concept of *meta-didactical conflict*, elaborated by Arzarello and Ferretti (2021). It exploits an analogy between the answers given by teachers in surveys like that illustrated here and the “incommensurable discourses” described by Sfard (2008). She shows how, in classroom interactions, interlocutors many times share the same words but with a different, incommensurable, meaning, of which they are not aware: a conflict is so generated. Something very similar we found when we analysed the answers to our questionnaire: here we found incommensurable languages. As the name we adopted suggests, this conflict consists of three components and is meta-didactic. We have defined it as meta-didactic because it is inherent in discourses about didactic processes such as assessment, students' skills and mistakes.

In this study we have highlighted how a twofold conflict emerges in many teachers' answers, because they interpret: (i) the difficulties of students in the INVALSI tests in a way that is completely different from what unquestionably appears from the data of the survey; (ii) the rationale of the INVALSI tests in a contradictory way (e.g., in the way how they couple the dyads suitable/not suitable Vs most used/not used in the examples with respect to what appears in the survey data).

To summarize our findings, we can say that teachers' answers to our survey show an overspread meta-didactic conflict. We consider worth of note, that even if teachers declare to consider the task highly suitable to assess students learning, they do not say to make an extensive use of this kind of task and, above all, in most of the cases, they are not always able to recognise all the suggested correct solution strategies. These observations can give a negative answer to our research question Q1. Furthermore, the case of the crochet placemats task is an example of how our survey resulted to be a useful tool in order to investigate teachers' interpretation of INVALSI data and their use in the teaching practices. In this sense, our result can give an answer also to research question Q2.

Conclusion

In the paper we have presented the first findings from a survey aimed at investigating the way primary school mathematics teachers read, interpret and use INVALSI Italian standardized assessment. According to the INVALSI theoretical framework (INVALSI, 2018), the crochet placemats task has been designed with the aim to evaluate if fifth grade students were able to solve a given problem of direct proportionality and to describe their solution strategies. At a meta level, in order to evaluate

the correctness of the solution to the task, teachers have been requested also to recognise the correctness of the strategies employed by the students. This needs a deep understanding of the underlying conceptual field of the multiplicative structures. Analysing teachers' responses to the survey, we have observed a basic meta-didactical conflict. The necessity of overcoming this conflict can suggest suitable designs for professional development programs for in-service teachers in order that they develop a correct culture of assessment and base their teaching practices on a correct interpretation of the data in national surveys. If we want to overcome the risk that standardised assessment is limited to classifying students, schools and nations, we need to develop the dialogue between standardised assessment and mathematics education. To fully recognise the potential and educational goals of standardised assessment, there is a need to develop and disseminate effective theoretical tools for interpreting the quantitative data and the macro phenomena emerged.

Moreover, our data provide several elements to be further analysed, both quantitatively and qualitatively, to shed light to critical aspects connected with teachers' perception of the usefulness of standardized assessment in their professional activities, that is the main general aim of our interdisciplinary research project.

Acknowledgment

The authors are grateful to the members of the research group: Ira Vannini, Giorgio Asquini, Barbara Balconi, Giorgio Bolondi, Elisa Guasconi, Violetta Lonati, Daniela Maccario, Annarita Monaco, Ottavio Rizzo, Elisa Truffelli and Valentina Vaccaro.

References

- Arzarello F. (2018). Insegnare, apprendere e valutare in matematica: un'interazione delicata, *L'insegnamento della matematica e delle scienze integrate*, 41A-B(2), 135–159.
- Arzarello, F., & Ferretti, F. (2021) The connection between the mathematics INVALSI test and the teaching practices: an explorative study. In Falzetti P. (Ed.). *I dati INVALSI: uno strumento per la ricerca*. Milano: FrancoAngeli.
- Bolondi, G., Ferretti, F., & Santi, G. (2019). National standardized tests database implemented as a research methodology in mathematics education. The case of algebraic powers. In Jankvist, U. T., den Heuvel-Panhuizen, V., & Veldhuis, M. (Eds.) *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education*. Freudenthal Group & Freudenthal Institute, Utrecht University and ERME, 4049–4056.
- De Lange, J. (2007). Large-scale assessment and mathematics education. *Second handbook of research on mathematics teaching and learning*, 2, 1111–1144.
- Di Martino, P., & Baccaglioni-Frank, A. (2017). Beyond performance results: Analyzing the informational and developmental potentials of standardized mathematics tests. *For the Learning of Mathematics*, 37(3), 39–44.
- Doig, B. (2006). Large-scale mathematics assessment: looking globally to act locally. *Assessment in Education: Principles, Policy & Practice*. 13(3), 265–288.

- Ferretti, F., Bolondi, G. (2021). This cannot be the result! The didactic phenomenon “the Age of the Earth”. *International Journal of Mathematical Education in Science and Technology*. 5(2).1464–5211. <https://doi.org/10.1080/0020739X.2019.1670366>
- Ferretti, F., Gambini, A. (2018), A vertical analysis of difficulties in mathematics by secondary school to level; some evidences stems from standardized assessment. In Dooley, T. & Gueudet, G. (Eds.) *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education (CERME10)*, Dublin, Ireland: DCU Institute of Education and ERME. 3492–3499.
- Ferretti, F. Gambini, A., & Santi, G. (2020). The Gestinv Database: A Tool for Enhancing Teachers Professional Development within a Community of Inquiry. In Borko, H., Potari, D. (Eds.) *Proceedings of the Twenty-fifth ICMI Study School Teachers of Mathematics Working and Learning in Collaborative Groups*; University of Lisbon: Lisbon, Portugal, 621–628.
- INVALSI (2018). *Quadro di riferimento delle prove di INVALSI di matematica*. Retrieved from https://invalsi-areaprove.cineca.it/docs/file/QdR_MATEMATICA.pdf
- Looney, J.W. (2011). Integrating Formative and Summative Assessment: Progress Toward a Seamless System? *OECD Education Working Papers*, 58, OECD Publishing.
- Meinck, S., Neuschmidt, O., & Taneva, M. (2017). Workshop Theme: Use of Educational Large-Scale Assessment Data for Research on Mathematics Didactics. In *Proceedings of the 13th International Congress on Mathematical Education*. Springer, Cham, 741–742.
- Sfard, A. (2008). *Thinking as communicating: human development, the growth of discourses, and mathematizing*. Cambridge University Press, Cambridge, UK.
- Suhr, D. (2006). Exploratory or Confirmatory Factor Analysis. *SAS Users Group International Conference*, Cary: SAS Institute, Inc., 1–17.