ROUND TABLE / TABLE RONDE

Uwe Gellert (chair), Freie Universität Berlin (Germany)

Gilles Aldon, ENS de Lyon (France)
Cristina Sabena, Università di Torino (Italy)

Lisa Björklund Boistrup, Stockholm University (Sweden)

Rossella Garuti INVALSI (Italy)
Francesca Martignone University of Eastern Piedmont (Italy)
Introduction to the Round Table on ‘Assessment in Mathematics Education: Resource or Obstacle’

Uwe Gellert, Freie Universität Berlin (Germany), uwe.gellert@fu-berlin.de

Abstract: In the format of a Round Table, we discuss the question of how assessment can support teaching and learning activities in the mathematics classroom from various perspectives. Our aim is to understand under which conditions and in which ways assessment is, or can be, a resource and not an obstacle for the teaching and learning of mathematics.

Résumé: Dans le format d'une Table Ronde, nous discutons de diverses perspectives la question de comment l'évaluation peut soutenir des activités d'enseignement et d'apprentissage dans la classe de mathématiques. Notre objectif est de comprendre dans quelles conditions et comment l'évaluation est, ou peut être, une ressource et non un obstacle pour l'enseignement et l'apprentissage des mathématiques.

Introduction

In 1993, the 45th conference of CIEAEM had been organised around the theme of Assessment Focussed on the Student. In its opening plenary, the at that time vice-president of CIEAEM, Lucia Grugnetti, stated:

The way mathematics instruction functions, as well as the entire spirit in which it takes place, is strongly influenced by assessment methods. Assessment is not just a separate appendix to mathematics instruction; it is one of its crucial components. (Grugnetti, 1994, p. 3)

Grugnetti’s view harmonises well with the ways in which many students experience the teaching and learning of mathematics in school. For many students the question whether assessment in mathematics education is a resource or rather an obstacle does not make much sense. Assessment in the mathematics classroom is so natural that school mathematics without assessment can hardly be imagined. However, as Clarke (1996, p. 327) points to, assessment is not a neutral element, but “a powerful mechanism for the social construction of mathematical competence”. Assessment serves the institutional functions of schooling of qualification, of cultural reproduction, and of allocation. It informs the students about the criteria for legitimate participation in social settings such as the classroom itself, but also in everyday contexts. It gives feedback on relative achievement in mathematics in the face of the expected learning outcomes; and it provides grounds for orienting the students towards the diverse vocational fields.

At the same conference in 1993, Leonor Cunha Leal and Paulo Abrantes distinguished four facets of assessment: (1) summative assessment, in which assessment is a measurement aiming at a score, (2) diagnostic assessment, the focus of which is on the preparedness of the student(s) for the coming mathematical topics, (3) formative assessment, focussing on the teacher’s control of the teaching-learning-process, and (4) assessment as a dynamic interpretation of student performance, in which problems are characterised and hypotheses generated. In the fourth sense, assessment is an “integral part of the learning process” (1994, p. 49). From the first towards the fourth facet, it seems to be obvious that assessment can increasingly be considered a resource for mathematical instruction in the classroom.

Cunha Leal and Abrantes’s clarification of the concept of assessment is located in the context of a developmental project in mathematics education. Beyond this particular context, and from more micro and more macro perspectives, further meanings of assessment can be identified. For instance, studies of mathematical classroom practice from a conversation analysis perspective clarify in which way assessment is an intrinsic element of the classroom discourse. Within the sequential structure of the classroom discourse, continuous assessment is obligatory for the discourse to
continue. If the teacher does not confirm the legitimate character of a student’s utterance, its conditional relevance remains open. It is the teacher’s reaction, which is based on her assessment of the pertinence of the student’s utterance, that decides whether the propositional content becomes included in the taken-as-shared knowledge of the class. That a teacher ‘assesses’ a student’s utterance by saying “yes”, “alright”, by nodding or else is first of all functional for the continuation of the classroom discourse. Only on a second plane, the micro assessment can also express praise (cf. Mehan, 1979; Streeck, 1979).

Another, more recent and more macro, aspect of assessment can be seen in the establishment of international comparative studies of student achievement and student competence which, originally, had been intended to measure the effectiveness and efficiency of national school systems (e.g., the PISA). Although it may be doubted if the methods for measuring achievement at system level can easily be transferred to the level of the learning individual, or if the test items used for the first purpose should have any influence on the arrangements for teaching and learning in the mathematics classroom, in many countries a considerable impact of PISA on mathematical education practices on the classroom level as well as on the curricular level has been noted (cf., Stacey & Turner, 2015). Policy makers and practitioner seem to experience this impact differently.

Against this background, the Round Table on assessment in mathematics education focuses on the question of how assessment, on its various levels, can support teaching and learning activities in the mathematics classroom. The aim is to discuss under which conditions and in which ways assessment is, or can be, a resource and not an obstacle for the teaching and learning of mathematics. The three contributions cover a broad spectrum of perspectives, some of them sketched above. Rossella Garuti and Francesca Martignone draw on their work in the context of the Italian assessment system. They discuss what seems to be necessary to convert the items and results of national standardised testing into resources for schools and teachers on the one hand, and for researchers on the other hand. Gilles Aldon and Cristina Sabena report on the European project Improving Progress through Formative Assessment in Science and Mathematics Education and its aim to use technology for formative assessment. They reflect on their classroom experience in contexts of low achieving students and show how formative assessment can serve as a didactical resource for teachers with the potential to foster meta-cognitive approaches to the learning of mathematics. Finally, Lisa Björklund Boistrup summarises her empirical classroom research with teachers in Swedish schools in which she identified four dominant assessment discourses of teachers. She discusses which of the four tend to build obstacles for students’ learning and which of them constitute resources. Such a categorisation can be useful for involving practitioners in reflections about possible effects of assessment practices.

The experience of the Round Table shows how fruitful it results to juxtapose perspectives on assessment in mathematics education that differ substantially in the educational strategies they facilitate, while pulling in the same direction: to make assessment a resource for mathematics teachers’ teaching and for students’ learning of mathematics.

REFERENCES


Actes de la 45ᵉ rencontre de la CIEAEM – Proceedings of the 45th CIEAEM meeting (pp. 3–9). Cagliari (Italy), 4–10 July 1993.


Formative assessment in the FaSMEd Project: reflections from classroom experiences

Gilles Aldon*, Cristina Sabena**

*ENS de Lyon
**Dipartimento di Filosofia e scienze dell’educazione, Università di Torino

Mathematical objects can be seen both as objects and as tools (Douady, 1986). For example, knowing a rotation is both understanding a formal definition of the transformation and its main properties of conservation of distances and angles, but also being able to use a rotation in solving a problem and/or in a proving process. Therefore, assessing the deep understanding of a mathematical object must take into account these different and complementary aspects.

But which kind of assessment do we speak of? The function of different assessments within an institution can be roughly split into two different roles: a role in certification of acquisitions of knowledge or competencies and a role in learning accompanying. The former refers to summative assessment and we will not discuss deeply this aspect of assessment here; the latter can be included in the process of formative assessment (FA) and appears to be a tool (a resource) for teachers in order to enhance mathematical students' learning. If teaching and learning are driven by summative assessment, the relationship to knowledge and to the learning process risks to be modified: instead of learning for understanding knowledge at stake, learning becomes understanding the way of succeeding in typical tests. Even if summative assessment and exams have a great importance in educational systems, this kind of assessment cannot be considered as a resource for acquiring knowledge. On the contrary, formative assessment can be seen as a resource for teachers and for students in the teaching and learning process.

Our contribution to this Round Table on assessment stems from our joint participation to the FaSMEd Project. FaSMEd (“Improving Progress through Formative Assessment in Science and Mathematics Education”) is a European Project (FPVII, 2013-15) aiming at investigating

the use of technology in formative assessment classroom practices in ways that allow teachers to respond to the emerging needs of low achieving learners in mathematics and science so that they are better motivated in their learning of these important subjects. (FaSMEd Project Document, p. 2)

Three main polarities can be identified within the project, considering the teaching-learning of mathematics and science: (i) formative assessment practices; (ii) the role of technology; (iii) attention to raise attainment, especially of low-achieving students.

Within FaSMEd, we share the definition or formative assessment given by Black & Wiliam (2009) on pragmatic basis, according to which

evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited (Black & Wiliam, 2009, p. 7).

This definition takes into account three key processes in learning and teaching (Ramaprasad, 1983):

- Establishing where the learners are in their learning;
• Establishing where they are going;
• Establishing what needs to be done to get them there.

It is also stressed that it is not only the teacher to be responsible for these processes, but the learners, both as individuals and as groups, play a crucial role as well.

Within this view, assessment is no longer considered as an object, but as a process that may change entirely the way teachers are organizing their lessons, and the way learners are managing their learning path.

The table of William & Thompson (2007; the table is in the annex) crosses the fundamental questions of formative assessment and the actors: the teacher, the class and the (generic) student. For example in the first cell of the table, the learning intentions and the criteria of success are pointed out. Overall, five key strategies are identified:

1. “Clarifying and sharing learning intentions and criteria for success;
2. Engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding;
3. Providing feedback that moves learners forward;
4. Activating students as instructional resources for one another; and
5. Activating students as the owners of their own learning”. (Black & William, p. 8)

Drawing on two examples from the classroom context, one in France and one in Italy¹, we will now point out the relationships between FA and the mathematical knowledge at stake, and some potentialities of new technologies in the teacher’s hand for formative assessment processes.

The first example (France) relates to a formative lesson about fractions. The teachers begins with a list of mathematical competencies the students have to acquire in the lesson:

• to read and to write a fraction,
• to code and to decode a fraction,
• to give equal fractions,
• to represent a fraction on a numeric line,
• to read a fraction represented on a numeric line,
• to compare a fraction to the unit and to another fraction with the same denominator.

The teacher proposes in the classroom a quiz focusing on these competencies and collects the results using a student response system. In the next lesson, the class results are given individually to the students: for each of these competencies, the students have the representations of their own capabilities using a representation of fractions that corresponds to the “Understanding learning intentions and criteria for success” of the table. The next lesson allows them to write themselves on the sheet their own result and to have a visual representation of their progression. For the teacher, the FA strategy corresponds to the “Engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding” because she organised the classroom depending on the success or the difficulty of each student relatively to the targeted competencies.

¹ The French team is formed by Gilles Aldon and Monica Panero. The Italian team is formed by Annalisa Cusi, Francesca Morselli and Cristina Sabena. See also the workshop by these authors in this volume.
Fig. 1. Representations of a student's capabilities about fractions, using fractions representations.

The second example (Italy) illustrates some potentialities of connected-classroom technologies for formative assessment, when suitably exploited by the teacher as a didactical resource. In a classroom where the students have difficulties in writing their reasoning (IV grade), the use of instant polls has worked as a means for supporting all the students to express their answer, and then explain them during a classroom discussion. For instance, faced with the students’ difficulties in correlating formulas to verbal expressions describing a given situation, the teacher shows two different representations (a word sentence like “adding always 5” and a formula, like “\(k=n\times 7\)”) and asks: what is written in the formula does correspond to what is written in the text? Three answers are given to choose from: yes, no (correct answer), and I don’t know.

Working in pairs with connected tablets, the students give their answers, which are then visualised on a whiteboard and showed through a bar diagram (Fig. 2). The grouped answers allow the teacher to get an immediate grasp on the overall situation of the classroom, so to realize that even with a task considered (by the teacher) relatively simple, some students either failed or did not dare to choose a yes/no answer. At the same time, the bar diagram allows the students to see the answers given by the other classmates, so to get a first feedback on one’s own answer (it is or it is not in the main stream).

Fig 2. The results of the instant poll on the meaning of text and formula, shared in a whiteboard so to foster discussion

By means of the class discussion (strategy 2, “Engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding”), the students are asked to express their reasoning behind their choice, and this constitutes a first reflective moment for them. Students
may also profit from hearing their mates arguments, and so compare their own reasoning to others’ (strategy 4, “Activating students as instructional resources for one another”).

These two short examples illustrate how it is possible to apply the FA framework to give a picture at a certain moment of a technology-based formative assessment lesson and to realize the dynamics that occurs when the teacher and the students become aware of their work.

The last cell of the table crossing the student's viewpoint and the questions “Where the learner is right now?” and “How to get there?” concerns the possibilities to activating students as the owners of their own learning, which can be related to an approach of meta cognition.

Acknowledgements

The research leading to this results has received funding from the European Community’s Seventh Framework Programme fp7/2007-2013 under grant agreement No [612337].

REFERENCES


Annex

<table>
<thead>
<tr>
<th>Role</th>
<th>Where the learner is going?</th>
<th>Where the learner is right now?</th>
<th>How to get there?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher</td>
<td>Clarifying learning intentions and criteria for success</td>
<td>Engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding</td>
<td>Providing feedback that moves learners forward</td>
</tr>
<tr>
<td>Peer</td>
<td>Understanding and sharing learning intentions and criteria for success</td>
<td>Activating students as instructional resources for one another</td>
<td></td>
</tr>
<tr>
<td>Learner</td>
<td>Understanding learning intentions and criteria for success</td>
<td>Activating students as the owners of their own learning</td>
<td></td>
</tr>
</tbody>
</table>

Aspects of Formative assessment (Wiliam & Thompson, 2007, as in Black & Wiliam, 2009, p. 8)
Assessment in mathematics education: Inevitable, but resource or obstacle? Different assessment discourses in mathematics

Lisa Björklund Boistrup, Department of Mathematics and Science Education, Stockholm University, S-106 91 Stockholm, Sweden, lisa.bjorklund@mnd.su.se

Abstract: In this paper, four assessment cultures (discourses) are presented where the assessment may provide with obstacles, or, resources, for the learning and engagement in mathematics. Assessment is in this text understood as broadly encompassing tests as well as teacher feedback in day-to-day mathematics teaching and learning. The four presented discourses were construed and adopted on a total in research in 30 mathematics classrooms with students in the ages of 7-16 years. The four discourses are «do it quick and do it right», «Anything goes», Openness to mathematics», and «Reasoning takes time». The first two assessment discourses hold various obstacles for students’ learning and engagement in mathematics, whereas the two last ones may constitute resources for students learning and active agency in mathematics. A claim made in the paper is that the four presented discourses offer teachers, students, and decision makers means to grasp essential aspects of assessment practices in mathematics classrooms including testing practices.

Résumé : Dans cet article, quatre cultures d'évaluation (les discours) sont présentés lorsque l'évaluation peut fournir des obstacles, ou, les ressources pour l'apprentissage et l'engagement en mathématiques. L'évaluation est dans ce texte comprend aussi les tests ainsi que des évaluations des enseignants dans l'enseignement et l'apprentissage des mathématiques au jour le jour. Les quatre discours présentés ont été interprété et adopté sur un total de recherche dans 30 classes de mathématiques avec des élèves dans les âges de 7 -16 ans. Les quatre discours sont «faire rapidement et le faire bien», «Tout va», «l'ouverture aux mathématiques », et« Raisonnement prend du temps ». Les deux premiers discours d'évaluation tiennent divers obstacles pour l'apprentissage et l'engagement des élèves en mathématiques, tandis que les deux derniers peuvent constituer des ressources pour les étudiants comme acteur en mathématiques. Les quatre discours présentés offrent aux enseignants, les étudiants et les décideurs les moyens de saisir les aspects essentiels de pratiques d'évaluation dans les cours de mathématiques, y compris les pratiques d'évaluation.

Assessment and feedback in mathematics classrooms

In this contribution to the round table at CIEAEM67 the focus is on assessment as a key notion relevant for problematizing all students’ possibilities to getting the opportunity to learn mathematics. When doing this I point at assessment cultures (discourses) where the assessment may provide with obstacles, or, resources, for the learning and engagement in mathematics. Assessment is in this text understood as broadly encompassing tests as well as teacher feedback in day-to-day mathematics teaching and learning (Björklund Boistrup, 2015b; Tunstall & Gipps, 1996). Feedback is then interpreted as conveying the teacher’s assessments to the student, through words (for example “Well done”) and/or body movements (for example a nod with a smile). Drawing on Ball et al (2012) and Foucault (2003; 2008) I discuss assessment as part of governings within the system of school. On one hand any mathematics classroom itself is immersed in a political context with decisions made on local and national levels, which have a direct, and governing, impact on the teaching and learning (Valero, 2004). On the other hand there are structural, often implicit, factors that contribute to how students are governed, and hence invited, or not invited, into mathematics (Straehler-Pohl & Gellert, 2013).

One example of a discussion of assessment and its consequences within the field of mathematics education is Morgan’s (2000) critique noting mainstream traditions of mathematics assessment research. Morgan emphasised research that adopts a social perspective, arguing that a main concern
of research from a social perspective is to understand how assessment works in mathematics classrooms and more broadly in education systems. As I see it, one consequence of this reasoning is that it is essential to discuss the effects of assessments within all kinds of mathematics education, both from the perspective of teacher-student communications but to also bring in the broader political context in the analysis. Today, many years after Morgan’s text, there still seem to be a rather modest interest in assessment in mathematics education research addressing policy and political matters as part of the research (Björklund Boistrup, 2015a). Words like assessment and evaluation may be mentioned but often these practices are taken for granted as a non-problematic part of education. In this paper I present assessment discourses which can constitute analytical tools for investigating how students are assessed, and invited – or not –, into mathematics education. The four presented discourses were firstly construed from five mathematics classrooms in grade four (Björklund Boistrup, 2010). After that they have been interpreted from more than 25 mathematics classrooms with students in the ages 7-16, and have been adopted in action research projects with mathematics teachers (Björklund Boistrup & Samuelsson, in preparation).

**Construal of assessment discourses in mathematics**

The term discourse is adopted drawing on Foucault (1993, 2003). A discourse is viewed as part of the institution of school. It goes beyond a particular mathematics classroom communication (taken here in a multimodal sense according to Van Leeuwen, 2005) and, consequently, the discourses presented in this text are possible to interpret from many mathematics classrooms. For the people who are part of a discursive practice, like teachers and students, the “rules” of the discourses affect how it is possible to act and what is possible to communicate (Foucault, 1993; 2003). An inspiration for adopting discourse as something smaller than entire disciplines is Walkerdine (1988) who construed a “testing discourse” where the teacher posed questions to which she already knew the answer.

All discourses were construed drawing on three initial analyses:

- Analysis of assessment acts (feedback) in mathematics classroom communication between teachers and students (described in Björklund Boistrup, 2010, chapter 5). In what direction is the feedback – from teacher to student, and/or vice versa? What directions – (dis)approving, (dis)agreeing/recognising, (dis)interest/(dis)engagement, checking, guiding, challenging – are mainly present?
- Analysis of the focuses of the feedback (Björklund Boistrup, 2010, chapter 6). Is it about the student as a person, no-mathematical procedures or mathematical processes? What processes are present, for example knowing mathematical facts, practicing/routine, reasoning/arguing, defining/describing, inquiring/problem-solving?
- Analysis of communicative resources including artefacts part of the assessment acts (Björklund Boistrup, 2010, chapter 7). What roles do different resources play in the assessment acts? How are communicative resources promoted or restricted? How are open questions and/or silences present in teacher-student communication?

**Four assessment discourses in mathematics education**

The first discourse, “Do it quick and do it right” has similarities to a traditional discourse of mathematics education described in the literature where the main “rule” is that the work should be done quickly and what is counted is whether an answer is right or not. The teacher’s feedback focuses on procedures with non or limited mathematical content. Feedback in this discourse typically focuses on whether an answer is mathematically correct or not, instead of why and how the answer may be counted as mathematically relevant. Another typical feedback focus concerns how many items from the textbook the student has accomplished. The affordances for students to be invited to learn mathematics are limited since they are not really invited to engage in any aspect of
mathematics through the feedback. Looking at the discourse from a multimodal approach, it may be possible to construe in writings when a teacher’s feedback on a test is focused on the number of correct answers, for example when a teacher writes 11/21 (11 points out of 21). Here it is important to keep in mind that the items on the test may well be mathematically rich and also inviting to the students. What is analysed here is mainly the subsequent feedback. In speech, teacher’s feedback where this discourse is construed can be really short, along with body movements, and describe whether the student’s work is correct or not or whether the student is doing the “right” thing. The affordances for students’ learning in this discourse are low.

The second discourse, “Anything goes”, is more of the opposite to the first discourse and a discourse where students’ performances, which can be regarded as mathematically inappropriate, are left unchallenged. There is not much articulated feedback apart from general approval. There is a presence of open questions, but challenges are not common. There are no critical discussions about students’ solutions, and wrong answers can be left unchallenged. The students are invited by the teacher to use whatever communicative resources they want, without any considerations by the teacher or the students on what resources that have most affordances for their learning at that specific occasion. Because the teacher values the students’ performance so often, the teacher, at the same time, takes the role as the main agent, as “the one that is evaluating”. Sometimes the teacher takes a more passive role in the discourse. S/he then does not interfere with students’ reasoning even though something wrong is demonstrated. The affordances for students’ learning in this discourse are low.

The third discourse, “Openness with mathematics”, has more of an open focus on mathematical processes. In this discourse the feedback goes both in the direction from teacher to student and vice versa. Occasionally, goals for the learning are present. Quite often the questions posed are open. The teacher and student often show interest in mathematical processes, and there is also an awareness of students’ alternative interpretations of tasks. Sometimes the student is challenged with respect to her/his continued learning. The focus is mostly on mathematical processes and sometimes on the student’s own reflection of her/his own learning. “Wrong” answers are here used as starting points for discussions, and it is always clear what can be considered mathematically correct. Various kinds of feedback from teacher to student are often communicated through questions. Different semiotic resources are acknowledged and at times the teacher promotes, whilst at other times restricts, the use of semiotic resources dependent upon the meaning making and learning process demonstrated by the student(s). This seems to be in order to serve the continuing learning process. The teacher and students communicate in longer utterances, but not more than a few utterances each time. In this discourse, there are considered to be affordances for students’ active agency and learning of mathematics.

Finally, the fourth discourse, “Reasoning takes time”, takes the characteristics of “Openness with mathematics” one step further with a slower pace and an emphasis on mathematics processes such as reasoning/arguing, inquiring/problem-solving, and defining/describing. In this discourse assessments take place in both directions between teacher and student. There are often instances of recognition of the students’ demonstrated knowing, which are sometimes in relation to stated goals, and the questions posed are mostly open ones. At times feedback as interest and engagement are communicated by the teacher to the student and vice versa. The students are often challenged towards new learning with the focus mainly on mathematical processes and the students’ reflections on her/his own learning. Here, most emphasis is on the processes inquiring/problem-solving, reasoning/arguing, defining/describing, and, occasionally, constructing/creating. Different communicative resources are acknowledged, and the use of semiotic resources can also be promoted or restricted when serving a certain process. In this discourse, silences in teacher-student interactions are common, and the possibility (for both teacher and student) to be silent serves the mathematics focus. Various kinds of feedback from teacher to student are often communicated, sometimes through open questions. Both the teacher and student can be active for longer periods of
time. In this discourse as well, the affordances for students to take active agency are high. The possibility to be quiet and think for a while promotes this potential agency. Similarly, the affordances for students’ learning of mathematics are high and include a wide range of mathematics processes.

**Concluding discussion**

To summarise, the first two assessment discourses hold various obstacles for students’ learning and engagement in mathematics, whereas the two last ones may constitute resources for students learning and active agency in mathematics. A point I want to make is that teaching without any assessment is not possible, since there always will be feedback taking place, and the kind of assessment that different students encounter may serve as a gatekeeper to the subject of mathematics.

The discourses presented here, with the connections between assessment acts, focuses in the mathematics classroom, and roles of communicative resources, offer teachers, students, and decision makers means to grasp essential aspects of assessment practices in mathematics classrooms including testing practices. There is positive power in an increased awareness of discourses like these, not the least since different students most likely encounter different assessment discourses in school, and hence different learning opportunity. For teachers, the discourses can be a starting point for identifying how various assessment discourse practices take place in the classroom. In such an activity, the implicitness of assessment practices is made more explicit. One example here is how the discourse “Do it quick and do it right” is possible to construe in the classroom, and possibly contrary to the teacher’s original plan. The reason for this can be governings within the system of school, for example, through a strong tradition within mathematics teaching and/or through demands from municipalities, where a dominant discourse such as “Do it quick and do it right” can be construed. Discourses like these can be a starting point for discussions about assessment practices and what kind of governings they hold among teachers and school heads, and among people responsible on the municipal and national level.

**REFERENCES**


Östlings Bokförlag Symposium.


The SNV (INVALSI) experience

Rossella Garuti* Francesca Martignone**
rossella.garuti@libero.it * INVALSI, Italy
francesca.martignone@uniupo.it **University of Eastern Piedmont, Italy

Abstract: In this paper we would like to show how the items and the results of the Italian national standardized tests can be considered a resource, not only for the Ministry, but also for schools, teachers and researchers. After a brief description of some peculiar aspects of the Italian Assessment System, we show an example of SNV items. Through the discussion of the item goals, we show how the standardized tests could be considered a resource for teachers and researchers.

Résumé: Dans cet article, nous voulons montrer comment les éléments et les résultats des tests standardisés nationaux italiens peuvent être considérés comme une ressource, non seulement pour le ministère, mais aussi pour les écoles, les enseignants et les chercheurs. Après une brève description de certains aspects particuliers du système d'évaluation italienne, nous montrons un exemple des éléments de la SNV. Grâce à la discussion des objectifs de l’exemple, nous montrons comment les tests standardisés pourraient être considérés comme une ressource pour les enseignants et les chercheurs.

The Italian assessment System (SNV)

The Italian Assessment System (SNV: Servizio Nazionale di Valutazione) started its work in 2008 through annual surveys conducted by the National Evaluation Institute for the School System (INVALSI) at different school grades. The INVAlSI develops standardised national tests to assess pupils' reading comprehension, grammatical knowledge and mathematics competency. Tests are administered at the end of the school year in grades 2-5-8-10. The results of a national sample are annually reported and they are public as well as the test items. From 2008, only for grade 8, the standardised SNV test is part of the national final examination, which is carried out at the end of middle school and it is organised by the school. Therefore, the SNV test contributes to the final assessment of the students.

SNV Framework

SNV investigations aim at taking a snapshot of schooling as a whole: in other words, it is an evaluation of the effectiveness of education provided by Italian schools. Currently, SNV tests are administered every year to all students in grade 2-5-8-10 (grade 6 was involved until 2010). The results of a national sample are annually reported and they are public as well as the test items. From 2008, only for grade 8, the standardised SNV test is part of the national final examination, which is carried out at the end of middle school and it is organised by the school. Therefore, the SNV test contributes to the final assessment of the students.

2 http://www.invalsi.it/areaprove/index.php
3 http://www.invalsi.it/invalsi/index.php
4 http://www.indicazioninazionali.it/
Another important reference is the UMI-CIIM curriculum "Mathematics for the citizen"\(^5\), which is based on results of mathematics educational research and has deeply influenced the last formulation of the National Guidelines.

The SNV Framework defines what type of mathematics is assessed with the SNV tests and how it is evaluated. It identifies two dimensions along which the questions are built:

- the mathematical content, divided into four major areas: Numbers, Space and Figures, Relations and Functions, Data and Forecasts;
- the link with the National Guidelines: each item is linked to a specific final objective.

This subdivision of content into four main areas is now shared at the international level: in PISA there are four content categories (Quantity, Space and Shape, Change and Relationships, Uncertainty and Data) and in TIMSS there are four content domains (Number, Geometry, Algebra, Data and Chance).

The SNV tests differ from PISA or TIMSS surveys not only for the frequency (annual vs. triennial), for the type of tested population (census vs. sample), for the chosen population (grade-based vs. age-based students for PISA), for the links with a National curriculum, but above all for the goals. As a matter of fact, the SNV test results aim at providing the Ministry with a national benchmark for the assessment of the Italian students at different grade levels, taking into account the national guidelines.

**An Example from the SNV test in grade 8**

The following example highlights the peculiar features of the items in the SNV tests.

<table>
<thead>
<tr>
<th>The teacher asks: &quot;Can an even number greater than 2 always be written as the sum of two different odd numbers?&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below are the answers of four students.</td>
</tr>
<tr>
<td>Who has given the correct answer, justifying it properly?</td>
</tr>
<tr>
<td>Antonio: Yes, because the sum of two odd numbers is an even number. (44.0%)</td>
</tr>
<tr>
<td>Barbara: No, because (6 = 4 + 2). (6.4%)</td>
</tr>
<tr>
<td>Carlo: Yes, because I can write it as the odd number that precedes it, plus 1. (34%)(^*)</td>
</tr>
<tr>
<td>Daniela: No, because every even number can be written as a sum of two equal numbers. (14.0%)</td>
</tr>
</tbody>
</table>

*correct

Figure 1: Item from SNV (2011-2012) in Grade 8 (indicating the percentage for each option)

The item in Figure 1 would like to assess a competence identified as a goal in the National Guidelines (i.e. to recognize and produce correct arguments based on theoretical knowledge). It arises in the context of the latest Italian research in mathematics education (Mariotti 2006; Boero et al., 2007). It somehow condenses the results of a wide research about the approach to argumentation and proof in mathematics, even with young students (Garuti & Boero, 1994). In this item, Grade 8 students are required to select arguments about the validity or non-validity of a statement: they must

choose the right answer with the correct justification. This item requires that the student understands that every even number can be written as \((2n - 1) + 1\). When it comes to number 2 the formula still holds, but the sum is between two equal odd numbers. The chosen options correspond to the more frequently observed behaviours of students. They all involve students’ understanding and exploration of the statement. In particular, option A, which had 44% of the answers, highlights a typical mistake. To answer the question it is not relevant that the sum of two odd numbers is always even. We consider questions of this type very important because:

- within a standardized test, they assess mathematical skills that are typical of the cultural aspect of mathematics (argumentation and proof);
- this kind of item shows the possibility of using algebra as a tool for supporting reasoning and consequently they push teachers towards a change of their practices as a result of the discussions they have in their schools about the nature of the SVN tests.

This type of item could be an important stimulus for teachers to reflect on, to consider a new approach to the culture of theorems at school, and to challenge standard teaching practices. Usually, in Italy (and possibly also in other countries) the teacher asks the students to understand and repeat proofs of statements rather than to produce conjectures themselves or to proof on their own. Arguing and proving activities are not generally common in Italy in the first years of secondary school (lower and upper), but we think that an early approach to theoretical thinking is important (Garuti & Boero, 1994). This item and its results may represent a kind of script for the construction of classroom activities on these aspects of mathematics education. In fact.

**SNV tests as a resource for teachers and researchers**

Being aware that the national assessment tests are a tool for the Ministry to give comparative information at the national level on students’ learning, in this Round Table we would like to highlight how the analysis of SNV tests can also be a resource for teachers and for researchers. The example analysed before could be an interesting subject of study and reflection for teachers and researchers also within teacher education programs. Regarding the research, we quote an example from the project "Ideas for the Research" funded by INVALSI. Starting in 2014 an educational study in mathematics on the SNV test was carried out. Some results of this study were presented at CERME9 (Branchetti et al, in press), at PME39 (Ferretti et al. 2014) and also at CIEAM (Lemmo et al, in these proceedings). The study proposes an integrated analysis, qualitative and quantitative, which can provide input for reflection on national assessment tests. The goal of the study is to build analytical tools to select “chains of items” (i.e. questions administered in successive levels that can be connected by qualitative and quantitative analysis) that could identify situations of difficulties related to specific topics. For example, the comparison and ordering of rational numbers, the management of their different representations, etc. The theoretical lenses developed in the first part of the research were shared with secondary school teachers. The focus was on structures and contents of the items and on the students’ possible answers.

In working with teachers, the researchers try to answer these questions: Which mathematical content knowledge is involved? What do the items assess? How do the students answer? Where do they make mistakes? Which could be the reasons for their behaviours?

This is an ongoing research, but it seems that the SNV items can become object of educational activities both in the classroom and in teacher education programs. It is clear that SNV tests cannot assess many important processes that can be mastered in a classroom by teachers, but teachers can use the information given by these standardized tests in order to identify possible students’ mistakes.

---

6 http://www.invalsi.it/invalsi/ri/sis/app_met.php
7 http://www.cerm9.org/
and misconceptions in tasks designed by taking into account the National Guidelines. The students’ possible strategies and mistakes can be object of discussion among teachers and researchers in a teacher education activity, arguing why and how these items and the answers may be mathematically relevant.

REFERENCES


