



**HAL**  
open science

# Mathematical written tests as formative assessment practice

Alice Lemmo

► **To cite this version:**

Alice Lemmo. Mathematical written tests as formative assessment practice. Thirteenth Congress of the European Society for Research in Mathematics Education (CERME13), Alfred Rényi Institute of Mathematics; Eötvös Loránd University of Budapest, Jul 2023, Budapest, Hungary. hal-04413550

**HAL Id: hal-04413550**

**<https://hal.science/hal-04413550>**

Submitted on 23 Jan 2024

**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.

# Mathematical written tests as formative assessment practice

Alice Lemmo

University of L'Aquila, Italy; [alice.lemmo@univaq.it](mailto:alice.lemmo@univaq.it)

*In this paper we present some considerations concerning mathematical written tests. They are often administered by teacher and educators with scoring purpose. Despite that, written mathematical tests belong to classroom assessment because they pursue curricular guidelines and support teaching and learning process all the way through school. In this study we present some suggestions for designing and assessing mathematical tests pursuing formative assessment. In particular, we propose suggestions for sharing data collected by such tests in order to provide both students and teachers with evidence which can be used and interpreted to promote and foster teaching and learning process.*

*Keywords: Summative assessment, formative assessment, written test.*

## Introduction

Popham (2000) considers assessment as the process “by which educators use students’ responses to specially created or naturally occurring stimuli in order to make inferences about students’ knowledge, skills, or affective status” (p.3). In other words, assessment include any activity in which evidence of learning is collected in a planned and systematic way and used to make decisions about learning. We choose a more specific definition of *classroom assessment* (cf. De Lange, 1999) as all the moment in which the teacher has as main purpose to use students’ results to get access to their skills and understanding; the aim is to improve further instruction for moving students’ learning forward. In both these definitions, assessment plays a crucial role in teaching and learning processes because it should provide feedback for all stakeholders in this process. In this prospective, one of the main purposes of research on classroom assessment is to use formative assessment techniques (see for example, Black & Wiliam, 2009) in teaching and learning practices.

In the last two decades, van den Heuvel-Panhuizen and Becker (2003) state that the use of tests in school is becoming more and more common in many countries over the years. This trend could be perceived as a remedy to the inconsistencies in classroom teachers’ assessment or grading practices (Shepard, 2000). Written tests are often conceived as objective measuring instruments; in contrast, individual or class discussions, systematic observations or other formative assessment practices can be conceived of as subjective. For van den Heuvel-Panhuizen and Becker (2003), this movement could represent the early beginnings of standardization in which tasks only measure the number of correct, missing and incorrect answers. This approach provides teachers with partial information on how to move teaching-learning process forward. In addition, it does not provide students with sufficient feedback to reflect on the tasks solving processes and how to improve their learning. Even assuming that classroom assessment is moving towards written test, it could continue to assume a formative value.

In this paper we present some considerations concerning mathematical written tests to explore whether and how they can be pursue formative assessment purposes (Black & William, 2009) both for students and teachers.

## Theoretical background

Written mathematical tests are frequently administered at the end of an activity or at the conclusion of a teaching and learning path before the beginning of a new one (van den Heuvel-Panhuizen & Becker, 2003). Actually, even if the teachers' aim is to certify the achievement of certain targets or goals and share it with external agents (for example parents), they belong to classroom assessment and so they should integrate with all teaching and learning practices with formative purpose. In this section we explore features of formative assessment with the aim of framing mathematical written tests as a form of its practice. Then, we also describe mathematical competence, assuming that assessment should weave together pedagogical and mathematics education research. We share Leder's (1992) view by which mathematics teaching, learning and assessment cannot be separated.

### Formative assessment

There are several definitions of formative assessment, in this contribution we choose the following:

“Practice in a classroom is formative to the extent that 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 and Wiliam, 2009, p.9).

Together with the definition, Black and Wiliam (2009) suggest five formative assessment strategies:

- a) clarifying and sharing learning intentions and criteria for success;
- b) engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding;
- c) Providing feedback that moves learners forward;
- d) Activating students as instructional resources for one another;
- e) Activating students as the owners of their own learning.

In previous studies, Black and colleagues (2003) proposed a further “sixth” strategy: *formative use of summative tests*. This was later ruled out because summative tests are considered as deficient of instant contingency that is perceived as a typical feature of formative assessment processes. However, we should highlight that Black and Wiliam (2009) suggest that ‘moments of contingency’ for the regulation of learning processes could be synchronous or asynchronous. Synchronous moments include immediate teachers' adjustments during one-on-one teaching or whole class discussion or activity; whereas asynchronous moments involve teachers' feedback through grading practices, and the use of evidence derived from data collected in a previous situation. In our opinion, written tests could allow to accomplish both moments in terms of the five formative assessment strategies (Black and Wiliam, 2009): for example, synchronous moments of sharing educational targets and criteria for success (a) and/or activating students as the owners of their learning (e); asynchronous moments after the returns of results and data collected through effective classroom discussions (b), providing feedback (c) and/or activating students as instructional resources for one another (d).

Another reason why they remove the “sixth” strategy is it doesn’t fit with the three key processes in learning and teaching (William & Thomson, 2007): where the learners are in their learning; where they are going and what needs to be done to get them there. Nevertheless, the authors claim:

“summative tests (or more properly, tests designed primarily to serve a summative function) provide ways of eliciting evidence of student achievement, and used appropriately, can prompt feedback that moves learning forward. These can also communicate to learners what is and is not valued in a particular discipline, thus communicating criteria for success. Where this has been done, it opens up the possibility of students helping one another, and using the tests as a guide to planning their own revision” (Black & Wiliam, 2009, p. 8).

Grabbing this last quotation and our previous considerations, we could hypotise that written tests could be tools for teacher to promote formative assessment.

### **Mathematical competence**

Assessment practice through tests impose to accurately define the expected results facing the educative actions (Pellerey, 2004). These results should be interpreted as “final behaviours that are observable and, to some extent, measurable” (Pellerey, 2004, p. 35, trad. Eng. by the author). Framing educational targets in terms of observable and measurable behaviour bears the risk of identifying the teaching-learning process with performances; that could be implemented in ‘teaching to the test’ or using Popham’s (2001) words in *item-teaching* instead of *curriculum-teaching*.

The quality of a performance does not depend only on the set of knowledge and skills that the performer may (or may not) have because these are grounded on several factors that cannot be directly observed. Therefore, a clear disciplinary framework for assessment is essential. In our opinion, this relates to the idea of competence as a complex and multidimensional process. A definition of mathematical competence could provide a tool for organising, framing, clarifying and addressing certain phenomena, situations or problems. Indeed, a model that describes mathematical competence could be useful in interpreting students’ difficulties, achieving a certain performance, creating suitable contexts for constructing, structuring and activating teaching interventions.

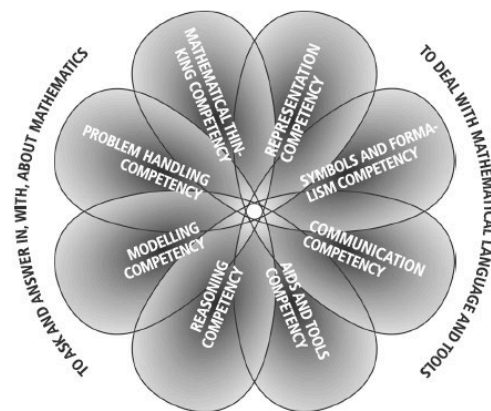
In selecting a theoretical model, we looked for a definition rich enough to be relevant in a complex situation, but at the same time aiming to reduce this complexity for all stakeholder’s usability (teachers, students, parents, ...). We choose the model of mathematical competency described by Niss and Højgaard (2019) because we share their perspective on mathematical competence:

“[...] the task we want to undertake [...] is to characterise what it means for someone to be mathematically competent. We want to do so in generic terms, i.e., in a manner that is independent of specific mathematical subject matter as well as of specific educational levels. This is a necessary prerequisite for coming to grips with what is common to mathematical competence across and beyond subject matter areas and across and beyond education segments, institutions and levels, as well as of fields of mathematical practice, and ultimately for coming to grips with what allows us to use the same term, mathematics [...]” (Niss & Højgaard, 2019, p.10).

The authors describe competence with a horizon point of view. This reflects our intentions to determine a model for assessment “across and beyond” subject matter areas and education segments.

In defining mathematical competency, Niss and Højgaard (2019) state that a primary goal of mathematical activity is to “pose and answer questions in or by means of mathematics” (p. 15). This ability has four different components: *fundamental mathematical thinking*; *posing and solving mathematical problems*; *dealing with mathematical models and modelling*; *undertaking mathematical reasoning*. Moreover, carrying out mathematical activity involves “handling mathematical language, constructs and tools” (p.17), that has again four different components: *dealing with mathematical representations*; *dealing with mathematical symbols and formalism*; *undertaking mathematical communication*; *dealing with material mathematical aids and tools*. These eight mathematical competencies constitute the conceptual model of mathematical competency.

In this conceptualization, competencies are of a purely cognitive nature. Even though the authors recognise that individuals’ affective, dispositional and volitional traits influence in mathematical activities, they are “multivariate functions of multitudes of background variables, life trajectory and experiences produced within and outside schooling and education, these traits are highly individualised.” (Niss & Højgaard, 2019, p. 18).



**Figure 1: Representation of the eight mathematical competencies (from Niss & Højgaard, 2019, p. 19)**

Figure 1 represents the eight mathematical competencies as a “flower”, in which each petal is different from but cross each other petal. These intersections are intentional because the competencies are not at all disjoint: each competency overlaps the other seven ones. Each competency has an own identity that distinguish it from the other. When someone use one of the them, some of the other may intervene in auxiliary roles.

We conclude the conceptualisation summary also recalling the three dimensions in an individual’s possession of a competency. The *degree of coverage* of a competency is the measure to which all aspects that frame the competency are part of the ownership of that competency. *The radius of action* of a competency represents the type of different circumstance and situations in which the individual can successfully activate the competency. Finally, the *technical level* indicates the level of sophistication of the mathematical concepts, results, theories and methods which the individual can bring to bear when exercising the competency.

The three dimensions suggest that defining whether or not an individual possesses one of the competencies is not possible; rather, a degree/level of competency can be identified depending on

how these dimensions intertwine. In this regard, the Italian National Guidelines for Assessment in primary school<sup>1</sup> identify 4 dimensions:

- *autonomy* in manifesting learning for a specific educational goal;
- *type of situation* (known/unknown) in which students reveal that he/she has achieved the goal;
- *resources* mobilised to solve the task; they could be presented by the teacher or spontaneously gathered in the learning context or previously acquired;
- *continuity in manifesting learning*. There is continuity when learning is enacted repeatedly or often. Alternatively, there is no continuity when learning is manifested sporadically or never.

These latter dimensions will be our benchmark to identify levels of each competency. The *degree of coverage* of a competency concerns all aspects of a competency possessed by the student and it could be related to the *resources* possessed or deployed by the learner, which may be their own or given by the teacher. The *radius of action* of a competency refers to the *type of situations* in which the individual successfully activates the competency. Finally, the *technical level* indicates the extent to which the student uses mathematical methods and can refer to his or her *autonomy* in dealing with situations and the *continuity* with which he or she manifests it.

Combining these dimensions results in different levels of competency (Table 1). Students with low levels of competency (1) have a narrow degree of coverage because they rely on the resources given by the teacher, a restricted radius of action referring to known situations and a low technical level that is expressed in a lack of autonomy and continuity. On the contrary, students with a high level of competency (4) demonstrate a broad radius of action in all situations, including unfamiliar ones; a wide degree of coverage that corresponds to the use of personal resources (individually elaborated); and finally a refined technical level managed with autonomy and continuity. Intermediate levels differ in the type of situations and resources. Level 2 students use teacher-provided resources in known situations and alternate continuity and autonomy; level 3 students handle unknown situations using both teacher and personal resources and alternate continuity and autonomy.

<i>Level</i>	<i>Autonomy</i>	<i>Type of situation</i>	<i>Resources</i>	<i>Continuity</i>
4	Yes	Known And Unknown	Personal and Given by the teacher	Yes
3	Yes	Known	Personal and Given by the teacher	Yes
	No	Unknown	Personal and Given by the teacher	Yes
2	Yes	Known	Given by the teacher	No
	No	Known	Given by the teacher	Yes
1	No	Known	Given by the teacher	No

**Table 1: Description of the levels from Italian National Guidelines for Assessment**

---

<sup>1</sup> <https://www.miur.gov.it/-/scuola-primaria-firmata-l-ordinanza-che-prevede-giudizi-descrittivi-al-posto-dei-voti-numeric-inviata-alle-scuole-insieme-ad-apposite-linee-guida>

## Suggestions for designing written mathematical tests

Very often, written mathematics tests are composed of tasks in which students are asked: to state definitions, theorems, etc.; to solve standardised exercises like those done during the school hours; to answer questions similar to those already covered in classroom practices and so on. This kind of test delivers a very dangerous message to both students and teachers that leads to the acquisition of an instrumental view of mathematics (Skemp, 1976). As van den Heuvel-Panhuizen and Becker (2003) propose, wide-ranging tasks allows the student to express him/herself and show with greater freedom what he/she knows and can do (see for example the flag problem presented in van den Heuvel-Panhuizen & Becker, 2003, p. 6-7).

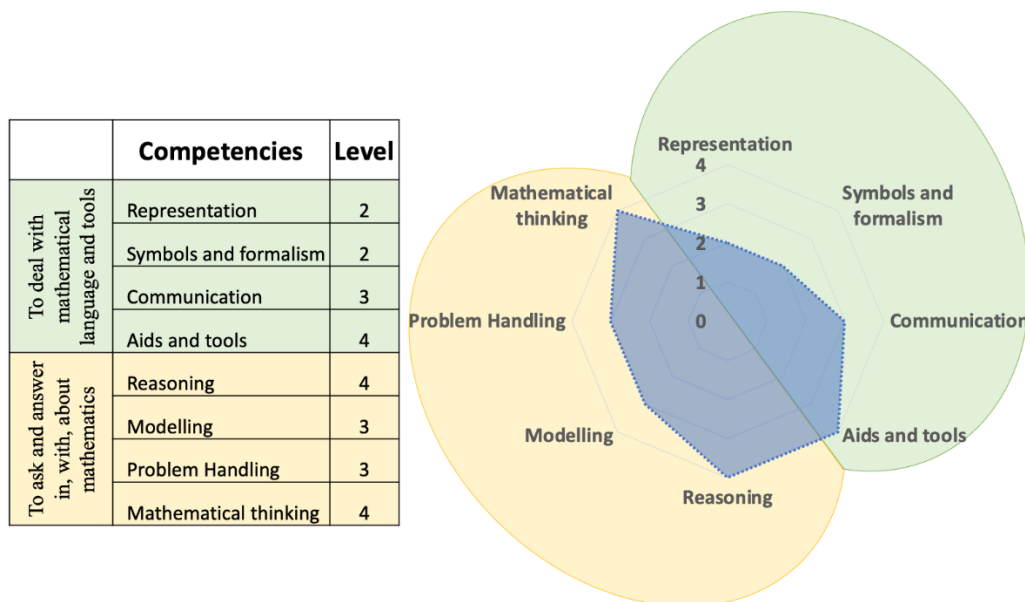
Designing a task that exclusively involves one of the different competencies is utopian, because each of them interacts with the others for the definition. For example, the simple request to determine the solutions of an equation calls for skills other than symbols and formalism competency. This is especially true with the wide-ranging tasks mentioned earlier. Beyond that, designing a single task relating to a single competency is not advisable: if for any reason a student did not respond to that task, the teacher would have no information about that specific competency. However, it is possible to identify tasks in which just a few are relevant for assessment. For example, if the task aims to assess *handling mathematical language competency*, teachers and students will focus on its four components and if there will be some errors or deficiencies in the other four, they could choose to ignore this. Share this assessment criteria allows teachers to clarify and share the assessment targets with students (a).

Administering tasks that focus on competencies is a way of giving students the opportunity to express themselves in mathematics. This aspect is also crucial on motivation. In a literature review, Harlen and Deakin Crick (2003) show that testing has a negative impact on motivation for learning that militates against preparation for lifelong learning. This reminds a second important formative assessment strategy: activating students as the owners of their own learning (e).

Feedback (c) plays an essential role in assessment; for this reason, designing the test composed by task that encourages the use of the eight competencies is not enough. It is important to set up the return of the data collected from the test. It could not be a single measure or mark because mathematical competency is a complex and multidimensional process. The return of data should therefore be equally complex and multidimensional in order to allow the student to grasp the multi-faceted nature of what he/she can do and knows. Only in this way he/she can use this data to make decisions about his/her own learning.

Using the four level defined in the previous paragraph for each competency allows teacher and student to observe and interpret what the student actually did in the test. This type of restitution also motivates the choice of wide-ranging tasks: it is not the solution the student finds that matters but the approach he/she takes to each task. In this way, the focus on the number or severity of errors and mistakes disappear and the attention is only in what student approach the tasks. This kind of approach require to observe the whole test and not a single task because each of the competencies lies in more than one task.

In Figure 2 we present what should be compiled by the teacher and returned to the student at the conclusion of the written test. The feedback the student receives in this case relates to the competencies in which he/she does or does not demonstrate in terms of autonomy, resources, situations, etc. In the example in Figure 2, the students show an advanced level in asking and answering in, with, about mathematics but he/she still has a basic level on dealing with mathematical language and tools. This return form can be a strong feedback on how to move forward with learning.



**Figure 2: Example of a student's return form concerning a written test**

During our teacher training courses, we ask secondary school teachers how they assess written mathematical test. Most of them state they use scores for determining the final mark. For example, some teachers choose a number of points that corresponds to each correct task, then they assess students' products counting the errors or missing answers they discover and subtracting part of the points from the sum of them. In this way, students' mark is based only on what they do not able to do and what is missing. On the contrary, the approach we present could allow teachers and students to go beyond the products and the number of right/wrong answers; it could let them to reflect on and analyse each competency and think about how moving learning forward.

We consider for example the follow exercise:

Find the domain and the zeros of the function:  $y = \frac{\sqrt{1-x} + 4}{\sqrt{3x-1}}$

For assessing this task within a mathematical written test, a teacher could choose that the correct answer of the task counts 5 points, and he/she could subtract one point for each error or 2 points if one of the two requirement is not satisfied (domain or zeros) or other. In the restitution of the whole test, the students will see the errors he/she did and/or the missing answers, the final mark and maybe some comments by the teacher. With our approach, students will continue to see the same aspects as before, but the mark will be different because the single task contributes for identifying the level for each competency. For example, even if the students gave a wrong answer, he/she could demonstrate that he/she is able to represent using some graphical signs or he/she could demonstrate high



communication competency if he/she correctly describes any solution passages and so on. In this case, the focus of the assessment is not on error or omission but only on the quality of the competencies the student employs regardless of whether the answer is correct or not.

We are aware that written mathematical tests are not enough for assessing mathematical competences, but this is not our goal. In this contribution, we would like to present and share an assessment approach that allows students and teachers to use test with formative and not standardised purpose.

## References

- Black, P., & Wiliam D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21, 5–31. <https://doi.org/10.1007/s11092-008-9068-5>
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Open University Press.
- De Lange, J. (1999). *Framework for classroom assessment in mathematics*. NICLA/WCER.
- Harlen, W., & Deakin Crick, R. (2003). Testing and motivation for learning. *Assessment in education: principles, policy & practice*, 10(2), 169–207.
- Leder, G. C. (1992). Curriculum planning+assessment=learning? In G. C. Leder (Ed.), *Assessment and learning of mathematics* (pp. 330–344). Australian Council for Educational Research.
- Niss, M., & Højgaard, T. (2019). Mathematical competencies revisited. *Educational Studies in Mathematics*, 102, 9–28.
- Pellerey, M. (2004). *Le competenze individuali e il Portfolio* [Individual competences and the portfolio]. La Nuova Italia.
- Popham, W. J. (2000). *Modern educational measurement: Practical guidelines for educational leaders*. Allyn and Bacon.
- Popham, W. J. (2001). Teaching to the test. *Educational Leadership*, 58(6), 16–20.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4–14.
- Skemp, R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching*, 77, 20–26.
- van den Heuvel-Panhuizen, M., & Becker, J. (2003). Towards a didactic model for assessment design in mathematics education. In A. J. Bishop, M. A. Clements, C. Keitel, J. Kilpatrick, & F. K. S. Leung (Eds.), *Second international handbook of mathematics education. Springer International Handbooks of Education, vol 10*. Springer.
- Wiliam, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53–82). Erlbaum.