Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education

Editors: Uffe Thomas Jankvist, Marja van den Heuvel-Panhuizen, Michiel Veldhuis
Organised by: Freudenthal Group & Freudenthal Institute, Utrecht University
Year: 2019

erme
european society for research in mathematics education

Editors

Uffe Thomas Jankvist, Danish School of Education, Aarhus University, Denmark.

Marja van den Heuvel-Panhuizen, Freudenthal Group & Freudenthal Institute, Utrecht University, Netherlands; Nord University, Norway.

Michiel Veldhuis, Freudenthal Group, Utrecht University, Netherlands; iPabo University of Applied Sciences, Amsterdam, Netherlands.

Publisher

Freudenthal Group & Freudenthal Institute, Utrecht University, Netherlands and ERME

ISBN 978-90-73346-75-8

Recommended citation for the proceedings

Jankvist, U. T., Van den Heuvel-Panhuizen, M., & Veldhuis, M. (Eds.). (2019). *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education*. Utrecht, the Netherlands: Freudenthal Group & Freudenthal Institute, Utrecht University and ERME.

Recommended citation for single entries in the proceedings

Authors (2019). Title of paper/poster. In U. T. Jankvist, M. van den Heuvel-Panhuizen, & M. Veldhuis (Eds.), *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education* (pp. xxxx-yyyy). Utrecht, the Netherlands: Freudenthal Group & Freudenthal Institute, University and ERME.

TABLE OF CONTENTS

Strategies of formative assessment enacted through automatic assessment in blended modality

Alice Barana[1] and Marina Marchisio[2]

[1]Department of Mathematics of the University of Turin; Turin; Italy; alice.barana@unito.it

[2]Department of Mathematics of the University of Turin; Turin; Italy; marina.marchisio@unito.it

This paper intends to contribute to the research on formative assessment in Mathematics providing a model of automatic assessment aimed at enhancing learning and self-regulation. The model was developed at the Department of Mathematics of the University of Turin (Italy). The main features of the model are: availability, algorithmic questions, open answers, immediate feedback, interactive feedback, and real-life contextualization. The effectiveness of the model to enact formative strategies is discussed though the results of a didactic experimentation involving 299 students of 8th grade, where automatically assessed assignments have been used both during Mathematics classes and as online homework.

Keywords: Automatic assessment, computer assisted instruction, feedback, formative assessment.

Introduction

It is widely acknowledged that assessment has a great influence on learning, impacting on when and how students work and learn. In particular, formative assessment practices help develop understanding and motivation, encouraging positive attitudes toward learning. Being responsive to the users' actions, digital technologies can make new room for formative assessment: with their capabilities of computing grades and offering feedback in real time, they can return information to students and teachers that is relevant to support and enhance learning processes. Web-based digital materials with automatic assessment can be used in face-to-face, blended or online courses; according to the modality adopted, they can facilitate personalized approaches as well as foster peer discussion.

This paper intends to contribute to the research on computer aided assessment in Mathematics, by proposing a model of automatic formative assessment and interactive feedback developed by the Department of Mathematics of the University of Turin (Italy). After a brief review of the literature on formative assessment, feedback and automatic assessment, the paper shows a model of automatic formative assessment using a system based on an Advanced Computing Environment, particularly effective for Mathematics. The model is discussed through some results of a didactic experimentation which involves 8th grade students.

Theoretical framework

Formative assessment

The term "formative evaluation" was coined by Michael Scriven in 1966 in opposition to "summative evaluation", to describe a practice aimed to collect information during a course in order to develop the curriculum (Scriven, 1966). Benjamin Bloom borrowed the term to indicate a strategy for mastery learning, namely a set of diagnostic-progress tests which should assess the achievement of the small units in which the program is divided (Bloom, 1968). Bloom's studies evidenced the effectiveness of this strategy to motivate students to forge ahead with the learning path. In 1989 D. Royce Sadler highlighted the role of feedback as a key distinction between formative and summative assessment. Sadler conceptualized formative assessment as the way learners use information from judgments about their work to improve their competence (Sadler, 1989).

More recently, Paul Black and Dylan Wiliam contributed to the development of a theoretical framework for the formative assessment, with particular reference to Mathematics Education. They spoke about formative practice, giving the well-known definition:

> 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 & Wiliam, 2009)

They individuated five key strategies through which formative practices can be enacted, involving students, peers and teachers: clarifying and sharing learning intentions and criteria for success; engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding; providing feedback that moves learners forward; activating students as instructional resources; and activating students as the owners of their own learning.

Feedback

The provision of feedback is the most distinctive feature of formative assessment. The power of feedback emerges in Hattie's metanalysis where it is considered one of the most effective strategies for learning (Hattie, 2009). John Hattie and Helen Timplerley expanded upon the model of good feedback, conceptualizing it as "information provided by an agent (e.g., teacher, peer, book, parent, self, experience) regarding aspects of one's performance or understanding" (Hattie & Timperley, 2007). Effective feedback, whose purpose is to reduce the discrepancy between current and desired understanding, indicates what the learning goals are, what progress is being made toward the goal and what activities need to be undertaken in order to make better progress. Feedback can work at four levels: the task level, giving information about how well the task has been accomplished; the process level, showing the main process needed to perform the task; the self-regulation level, activating metacognitive process; the self-level, adding personal evaluations about the learner.

Sadler emphasizes the focus on the learner's processing of feedback, noticing that if the information is not elaborated by the learner to alter the gap between current and reference performance, it will not have any effect on learning (Sadler, 1989). In order for feedback to be effective, students have to possess a concept of the standard being aimed for, compare the actual level of performance with the standard and engage in appropriate action, which leads to some closure of the gap.

Besides improving understanding, feedback can also be effective to enhance self-regulation, a process whereby learners set goals for their learning and monitor and regulate their cognition, motivation and behavior through internal feedback (Pintrich & Zusho, 2007). The process of generation of internal feedback can be facilitated by well-designed feedback which, according to Nicol and Macfarlane-Dick (2006), should clarify what good performance is; facilitate the development of self-assessment; deliver high quality information to students about their learning; encourage teacher and peer dialogue around learning; encourage positive attitudes, motivation and self-esteem; provide opportunities to close the gap between current and desired performance; and provide information to teachers that can be used to help shape the teaching.

Automatic assessment

Automatic assessment is one particular form of Computer Aided Assessment, characterized by the automated elaboration of students' answers and provision of feedback. Multiple choice is the most common question format; it is supported by the majority of online platforms, even though it considerably limits the cognitive processes involved in answering, especially in Mathematics (Bennett, 2012). To overcome this limitation, research centers and universities started to develop systems that are able to process open-ended answers from a mathematical point of view and to establish if they are equivalent to the correct solutions. Examples of similar Automatic Assessment Systems (AAS) are STACK, relying on the Computer Algebra System (CAS) Maxima (Sangwin, 2015) and Maple T.A., running on the engine of the Advanced Computing Environment (ACE) Maple (Barana, Marchisio, & Rabellino, 2015). By exploiting programming languages or mathematical packages, these AASs allow to build interactive worksheets based on algorithms where answers, feedback and values are calculated over random parameters and can be shown in different representational registers. Thus, new solutions for computer-based items can be conceived, including dynamic explorations, animations, symbolic manipulation that offer students experiences of mathematical construction and conceptual understanding (Stacey & Wiliam, 2013).

A model of formative automatic assessment for Mathematics

The Department of Mathematics of the University of Turin has designed a model for creating questions conceived for the formative assessment of Mathematics, using Moebius Assessment. (formerly known as Maple T.A.). This AAS was chosen for its powerful grading capabilities, for the robust mathematical engine running behind the system and for the possibility of integration within the Virtual Learning Environment (VLE) Moodle (Barana & Marchisio, 2016).

Aiming at enhancing learning and self-regulation, automatically assessed assignments should have the following features (Barana, Conte, Fioravera, Marchisio, & Rabellino, 2018):

- availability: students can attempt the assignments, integrated in a VLE, at their own pace, without limitation in data, time and number of attempts;
- algorithm-based questions and answers: random values, parameters or formulas make questions, and their answers, randomly change at every attempt. Though variables based on algorithms, random parameters, mathematical formulas, graphics and even animated plots can be shown in questions and feedback;
- open-ended answers: the multiple-choice modality is avoided whenever possible; open answers, given in different representational registers (words, numbers, symbols, tables, graphics) are graded through algorithms which verify if the student's answer matches the correct one, independently of the form;
- immediate feedback: results are computed in a very brief time and they are shown to the students while they are still focused on the task. Brief assignments are advised, in order to increase the immediacy of feedback;

- interactive feedback: just after giving an incorrect answer, the system can go through a step-by-step guided resolution that interactively shows a possible process for solving the task, which recalls previous knowledge and engage students in simpler tasks. They can gradually acquire the background and the process that enables them to answer the initial problem. They earn partial credits for the correctness of their answer in the step-by-step process;
- real-life contextualization: whenever possible, questions refer to real-world issues, which contributes to the creation of meanings and to a deeper understanding, as students can associate abstract concepts to real-life or concrete objects.

The feedback provided through this model acts not only at the task level, giving information about the correctness, but also at the process level, showing the steps toward the solution. Moreover, it can also act at the metacognitive level, providing opportunities for self-assessment, engaging less motivated students in active drills and offering partial credits for correct answers. This feedback satisfies the conditions individuated by Nicol and Macfarlane-Dick (2006) for the development of self-regulation. This model is particularly relevant in making students elaborate the feedback, as it is displayed interactively while they are still engaged with the task. Feedback can be effective according to Sadler's model: in fact, the interactive feedback offers a concept of standard that students can actively possess; immediate feedback helps them compare the actual level of performance with the standard; when trying the assignment again, students find similar tasks with different values, so that they are engaged in an activity that makes them repeat the process until mastered.

Research questions

The focus of this paper is to show how automatic assessment, implemented in a blended modality through classroom activities and online work, allows the enactment of formative strategies in order to enhance learning and self-regulation. In particular, the paper investigates whether the interactive feedback can be effective according to Sadler's model and whether the blended use of the automatic assessment can support formative assessment's strategies.

Didactic experimentation and data collection

A set of materials designed according to the model illustrated above has been proposed to 13 8th grade classes of 6 different lower secondary schools in the town of Turin, for a total of 299 students, during the school year 2017/2018. Interactive materials with automatic assessment, organized in 10 different units, were created by university experts and inserted in a dedicated VLE. The tasks were mainly designed using items form the INVALSI surveys, the national standardized tests that take place annually in Italy, in collaboration with an INVALSI expert, expanded and adapted to the automatic assessment. Mathematics teachers, working in close connection with the researchers, could use the materials in two modalities:

- in the classroom, with the support of the Interactive Whiteboard (IWB), where tasks were displayed. Students, in small groups of 3 or 4, were asked to solve one task. All the answers were collected by the teacher, one was collectively selected to be checked using the AAS. After verifying whether it was correct, all the groups, in turn, had to show their solving process to the others. The interactive worksheets displayed at the IWB supported the collective discussion and gave prompts for deeper reflection.

- for homework, using the online assessment and the interactive feedback to check understanding. Students could autonomously navigate within the platform and make one or more attempts to the assignments.

One example of question is shown in Figure 1. It asks students to solve a problem about linear models, open to different approaches. Students who give the incorrect answer to the main task are engaged with the exploration of the situation through a table to complete and a graphic to draw interactively. This question has been used online with some classes and in the classroom with others. On the platform, students were guided through the solving process and could repeat the problem with different data; in classroom, only the main task was displayed initially and the different solutions made by the groups of students were shared and discussed, with the support of the automatic assessment. During the experimentation, a PhD student participated to the lessons regarding the module on functions and modeling, the target topic of the experimentation. Lessons were videotaped and all data from platform usage were extracted in order to study the use of the assignments. The appreciation of the activities was measured though a questionnaire distributed at the end of the school year.

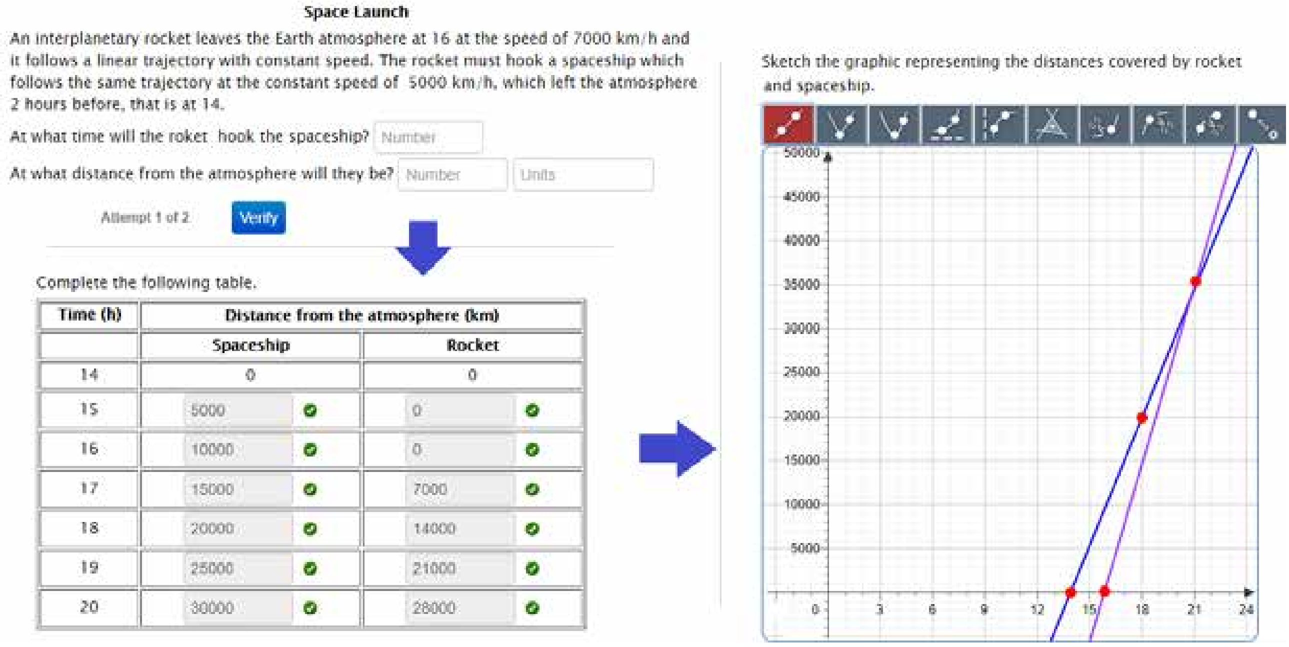

Figure 1: Example of automatically assessed question

Results and discussion

In order to evaluate whether the interactive feedback was effective according to Sadler's model, the usage and results of online assignments were analyzed. For each student the average number of attempts per assignment has been computed: it ranges from 1 to 12, with an average value of 1.70, which expresses students' tendency to repeat the questions. Then, the average grade each student earned in their first assignment attempt was compared with the average grade they earned in their last attempt through a pairwise Student t-test: the mean of initial grades resulted to be 51.55/100 (standard deviation: 19.63), while the mean of final grades was 59.02/100 (standard deviation: 19.87); the increase is statistically significative (p <0.0005). These results show that students used the information provided in the feedback to improve their results in subsequent activities.

Answers to the final questionnaire show students' perceived effectiveness of the automatic assessment. Table 2 reports the percentage of students' answers, given in a Likert scale from 1 (completely disagree) to 5 (completely agree). In particular, it emerges that students appreciated the usefulness of the group work in the classroom, supported by the use of the platform, and the online tests with immediate feedback. The low percentages of negative answers show that the feedback obtained through automatic assessment and peer work was useful to develop conceptual knowledge, since students could better understand the topics; it was also effective at a process level, to understand how to solve the problems; and at a metacognitive level, developing awareness of one's preparation.

Questions	**1**	**2**	**3**	**4**	**5**
Working in group was useful.	1.4%	6.3%	23.1%	38.5%	30.8%
Class activities were useful to better understand some Mathematical topics.	1.4%	4.5%	29.1%	45.9%	19.1%
Using the platform during classroom lessons was useful.	0.5%	3.6%	32.1%	47.1%	16.7%
I appreciated the possibility to revise the material used in class through the platform.	0.0%	8.2%	35.3%	36.4%	20.0%
The online exercises were useful to better understand the topics.	2.5%	5.6%	20.5%	46.0%	25.5%
The online exercises helped me make clear if I understood the topics.	2.5%	14.3%	26.7%	37.3%	19.3%
It is useful to have the correct answer displayed immediately after answering.	1.2%	3.7%	13.0%	36.0%	46.0%
The immediate assessment helped me understand how to answer the questions.	1.2%	5.0%	19.9%	32.3%	41.6%
Problems with step-by-step guided resolution helped me understand how to solve the exercises.	1.2%	8.7%	26.7%	34.2%	29.7%
Online exercises helped me to be aware of my preparation.	3.7%	8.1%	28.6%	40.4%	19.3%

Table 2: Student's answers to the final questionnaire, given in Likert scale 1 to 5

The open answers to the questions "what did you appreciate the most?" and "what were the online assessments useful for?", evidenced how the use of automatic assessment both in classroom and online supported the enactment of the strategies of formative assessment. Comments such as: "in my opinion, online exercises are useful to better understand the process to build a formula", "the platform was very useful to better understand both the current topic and the resolution of problems" and "I think that the online tests were useful, because they showed me the many possibilities to solve one problem" show how online assignments could clarify and share criteria for success. Tasks in the online tests were effective to elicit evidence of students' understanding, as other students report: "online tests were useful because only by doing them I was able to acknowledge whether I had understood a topic or not", "online tests were useful to individuate the points where I should improve". Other comments prove that automatic and interactive feedback supported learning improvements: "I could have immediate access to the result to understand where I made a mistake and I had the chance to make another attempt with different data to drill", "with the book you can do the assignments but you can't acknowledge if your resolution is correct, while the platform always shows you both the correct answers and the solving processes", "if you don't understand a topic, with guided exercises you can gradually learn how to solve them autonomously", "in classroom it often happens that I think I have understood a topic, but the next day I can't understand it anymore, while with the platform I can work whenever I need to". The assignments were also effective to activate students as owners of their own learning, as they themselves report: "technology encourages youngsters to work hard with homework, therefore they can better understand Mathematics", "during classwork or homework I often felt willing and happy to solve the exercises, and they were useful to understand the lesson", "I appreciated the online tests because they made me reason and work hard; sometimes I also had fun when doing Mathematics". Groupwork was very appreciated and it allowed students to be activated as learning resources one for another, as their comments show: "classroom lessons were useful because, besides solving problems, we had to interact with each other and this allowed us to better understand the tasks", "I appreciated reasoning together on the things that we were not able to do", "I appreciated sharing ideas with classmates and helping each other".

Conclusions

The results of the experimentation reported above show that the blended use of the online assessment made it possible to activate all the agents of the formative assessment (students, peers and teacher) and all the 5 key strategies individuated by Black and Wiliam: in the classroom, students received feedback from discussion and sharing ideas with peers, while on the platform the interactive feedback offered a guided support for understanding concepts and processes. This conception of automatic assessment provides enhancements with respect to paper-and-pen work and traditional book exercises: students can individuate their mistakes and make more attempts to improve their understanding, they can be actively engaged in Mathematical work and even have fun, although items are not game based. The results of this experimentation gave prompts for the activation of other research projects, aimed at studying the impact of these methodologies on students of different levels and backgrounds, as well as for the application of this model of formative automatic assessment to the learning of other subjects, even outside the STEM area, thus starting interdisciplinary collaborations.

Acknowledgments

We would like to thank Stefano Boffo, Francesco Gagliardi (CNR) and Rodolfo Zich (Fondazione Torino Wireless) for promoting the project and for participating to its realization. We are also grateful to Rossella Garuti (INVALSI) who collaborated to the development of the didactic materials.

References

Barana, A., Conte, A., Fioravera, M., Marchisio, M., & Rabellino, S. (2018). A Model of Formative Automatic Assessment and Interactive Feedback for STEM. In *2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)* (pp. 1016–1025). Tokyo, Japan.

Barana, A., & Marchisio, M. (2016). Ten Good Reasons to Adopt an Automated Formative Assessment Model for Learning and Teaching Mathematics and Scientific Disciplines. *Procedia - Social and Behavioral Sciences*, *228*, 608–613. https://doi.org/10.1016/j.sbspro.2016.07.093

Barana, A., Marchisio, M., & Rabellino, S. (2015). Automated Assessment in Mathematics. In *2015 IEEE 39th Annual Computer Software and Applications Conference* (pp. 670–671). Taichung, Taiwan: IEEE. https://doi.org/10.1109/COMPSAC.2015.105

Bennett, R. E. (2012). Inexorable and Inevitable: The Continuing Story of Technology and Assessment. *The Journal of Technology, Learning, and Assessment*, *1*(1), 1–22.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, *21*(1), 5–31.

Bloom, B.S. (1968). Learning for Mastery. *Evaluation Comment*, *1*(2).

Hattie, J.A.C. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. Routledge.

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Nicol, D.J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and 7 principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199-218.

Pintrich, P.R., & Zusho, A. (2007). Student motivation and self-regulated learning in the college classroom. In R.P. Perry & J.C. Smart (Eds.), *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Perspective* (pp. 1-8). Springer, Dordrecht.

Sadler, D.R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119–144.

Sangwin, C. (2015). Computer Aided Assessment of Mathematics Using STACK. In S.J. Cho (Ed.), *Selected Regular Lectures from the 12th International Congress on Mathematical Education* (pp. 695–713). Cham: Springer International Publishing.

Scriven, M. (1966). *The methodology of evaluation*. Lafayette, Ind: Purdue University.

Stacey, K., & Wiliam, D. (2013). Technology and Assessment in Mathematics. In M.A. Clements (Ed.), *Third International Handbook of Mathematics Education* (Vol. 27, pp. 721–751). New York: Springer Science+Business Media.