



Advantages of Using Automatic Formative Assessment for Learning Mathematics

Alice Barana¹ , Marina Marchisio² , and Matteo Sacchet¹ 

¹ Department of Mathematics, Università di Torino, Turin, TO, Italy
{alice.barana,matteo.sacchet}@unito.it

² Department of Molecular Biotechnology and Health Sciences,
Università di Torino, Turin, TO, Italy
marina.marchisio@unito.it

Abstract. Automatic Assessment Systems empowered by mathematical engines allow the development of online assignments for Mathematics, which goes beyond multiple-choice modality. Automatically assessed assignments, used with formative purposes, can support teaching and learning from several perspectives, such as conceptual and procedural understanding, metacognition, enactment of adaptive strategies, and teachers' management of the class. This paper reports on an experimentation where automatic assessment has been used in a blended modality according to a model of formative assessment and interactive feedback to enhance learning. The experiment involved a total number of 546 students of 8th grade in the town of Turin (Italy). The use of the automatic assessment is shown and exemplified. Data from learning tests, questionnaire and platform usage are analyzed and used to show the effectiveness of the interactive materials for enhancing mathematical understanding and self-assessment skills. Moreover, a profile of the students who did not use the online opportunities, defined as "reluctant users", is drawn and discussed.

Keywords: Automatic assessment · Formative assessment · Mathematics education · Reluctant users · Self-assessment

1 Introduction

In the last decades, national and international directives have largely encouraged the adoption of learning technologies at the school level, prompted by the idea that a stronger technological education will help the future citizens to cope with a digitalized society, to find employment opportunities in a revolutionized industry and to understand the continuous developments that affect our world [1]. Moreover, there are several findings in the research which show that technologies can generate positive effects on students, both from a cognitive and from a metacognitive point of view [2].

One of the features of technology identified as a key promoter of learning, especially in the research in Mathematics Education, is its interactive nature: stimulated by the student's action, a digital tool reacts and returns some information [3]. This kind of feedback, when processed by the learner, can be relevant for several issues, such as fostering deep reflection, putting the learner at the center of the learning process and

helping students become responsible for their learning [4]. Under this perspective, the automatization of feedback can represent a considerable advantage and digital learning materials with automatic assessment can be valuable resources, since they are easily sharable by teachers, accessible by students and they offer interesting learning experiences to improve understanding.

On the students' access to technologies, some issues can be raised that potentially hinder the widespread adoption of digital tools in education. First of all, the equity issue must be taken into account: digital tools should support every student's opportunity to learn fundamental Mathematics, as well as other subjects. Equity mainly involves socio-economic circumstances, but also differences in physical capabilities, gender or teachers' development should not prevent students from getting quality instruction [5]. Secondly, the complexity of the technological tools should be suitable to the students' digital skills: some findings suggest that "digital native" students do not easily transfer their skills in browsing social networks to a fluent use of technologies in a learning context [6]. Lastly, while the use of technology at school is commonly appreciated by the students, there is still a small percentage of pupils that prefer a paper-and-pen approach to learning. The little appreciation of computer-aided instruction seems to be correlated to the dislike of computers in general [7] or to the little perceived usefulness of the learning activities with that particular instrument [8].

This paper deals with the use of automatic assessment to design digital materials for learning Mathematics at lower secondary school. After an overview of the state of the art in the field of digital and automatic assessment in Mathematics, an experimentation on the use of automatic formative assessment in a blended Mathematics course for students of grade 8 is presented. The use of digital assessment by students is analyzed in light of the results of learning tests and of appreciation questionnaires, to the purpose of comparing the results of the students who did use the technologies with those of the students who made little use of the online activities. An analysis of the group of the "reluctant users" of the technology is made in order to understand the reasons for their behavior.

2 Theoretical Framework

2.1 Automatic Formative Assessment

Automatic formative assessment (AFA) in Mathematics is widely used in online courses, with the aim of keeping the learner involved and increasing motivation. Research centers and universities have developed 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, in order to widen the possibilities of student interaction with the digital tools. Example of similar Automatic Assessment Systems (AAS) are CALM, developed by the Heriot-Watt University in the programming language of Pascal [9]; STACK, developed by the University of Edinburgh and relying on the Computer Algebra System (CAS) Maxima for its Mathematics capabilities [10]; Moebius Assessment, developed by the University of Waterloo and running on the engine of the Advanced Computing Environment (ACE) Maple [11]. By Exploiting

programming languages or mathematical packages, these AASs allow instructors to build interactive worksheets based on algorithms where answers, feedbacks and values are computed over random parameters and can be shown in different representational registers. The use of similar AASs becomes powerful when combined with the principles of formative assessment. According to Black and Wiliam [12], “a practice 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”. Among the strategies for enhancing formative assessment, the provision of feedback that moves learners forward is one of the most acknowledged. Hattie and Timperley’s [13] definition of feedback conceptualizes it as “information provided by an agent regarding aspects of one’s performance or understanding”; they draw a model for designing effective feedback, which should provide relevant information about learning achievements, learning goals and current performance.

Research and experimentations have shown that digital materials which provide a feedback through an AAS empowered by a mathematical engine can be useful under different perspectives, such as learning, problem solving, metacognition, adaptive teaching and teacher’s practice.

First of all, they are effective to *enhance learning*. Taking advantage of all the potential of the technology, questions can now be created that could not be replicated with paper and pen, but that stimulate students’ cognitive processes. Stacey and Wiliam [5] suggest new solutions for computer-based items where situations or solutions can be explored dynamically; Sangwin [14] shows examples of implicit feedback provided to the student instead of explicit solutions and in relation to their answer, in order to enhance reflection and promote understanding. Questions can be enriched with dynamic images, animations, geometrical visualizations, symbolic manipulations, tables and other features that can be created through the computing environments running behind the ACE, thus offering students experiences of mathematical construction and conceptual understanding [15].

Secondly, they help *master problem solving procedures and strategies*. Beevers and Paterson discuss the use of step-by-step automatically graded resolutions of complex tasks that students are not able to solve on their own [16]. They state that the stepped approach guides learners towards breaking problems into smaller and manageable parts and helps them acquire control over the solving process. Students can thus develop procedural knowledge that, in Mathematics, is strictly connected to conceptual understanding [17]. Adaptive stepping capabilities are supported by several AASs and this area of research flows into studies on adaptive tutoring systems, which are at the most cutting-edge solutions in terms of e-assessment [18].

Third, automatic assessment proves to be useful at a *metacognitive level*. Nicol and Milligan [19] have analyzed e-assessment in the light of seven principles of good feedback practices to promote self-regulation, suggesting several question formats or uses of the online assessment in order to help students become aware of their knowledge, draw abstract concepts out of the examples, reflect on their own mistakes, and set goals for their learning.

Another relevant opportunity offered by interactive materials with AFA is that of *facilitating adaptive teaching strategies*. In fact, formative automatic feedback can work as an online tutor while students attempt the assignments at their own pace; the absence of restriction on the number of attempts allows weaker students to repeat the questions, while their most skillful classmates can put themselves to test with new ones [20]. Data collected by the platform can help teachers monitor learning and adjust instructional strategies accordingly. Moreover, digital interactive materials are particularly helpful for students with learning disabilities: in fact, the organization and presentation of the worksheets, the multimodality and the computing environments enhanced by technology can make up for their cognitive difficulties and motivate them to study.

Lastly, *teachers' work* can benefit from AFA, and not only for the time saved from manual correction. As mentioned above, data collected from the platform inform them about students' understanding and they can be useful to shape teaching and to have a complete picture of the students' gains. Many systems also allow teachers to author their own materials, or to adjust existing ones, and to share their work with colleagues. This permits to overcome dissatisfaction with textbooks and contributes to the development of professional communities and the growth of teachers' competences. Moreover, assignments prepared for one class can be easily reused and transferred to new courses the following years, thus the effort for producing online tests yields advantages and time saved in the future [21].

Using Maple TA AAS, the Department of Mathematics of the University of Turin has designed a model for the automatic formative assessment for STEM (Science, Technology, Engineering and Mathematics) disciplines, based on the following principles [22]:

1. availability of the assignments to the students, who can work at their own pace;
2. algorithm-based questions and answers, so that at every attempt the students are expected to repeat solving processes on different values;
3. open-ended answers, going beyond the multiple-choice modality;
4. immediate feedback, returned to the students at a moment that is useful to identify and correct mistakes;
5. contextualization of problems in the real world, to make tasks relevant to students;
6. interactive feedback, which appears when students give the wrong answer to a problem. It has the form of a step-by step guided resolution, which interactively shows a possible process for solving the task.

The last point recalls Beevers and Paterson's step-by-step approach to problem solving with automatic assessment, but here it is conceptualized in terms of feedback, highlighting the formative function that the sub-questions fulfill for a student who failed the main task. The interactive nature of this feedback and its immediacy prevent students from not processing it, a risk well-known in literature, which causes formative feedback to lose all of its powerful effects [23]. Moreover, students are rewarded with partial grading, which improves motivation [24].

2.2 Reluctant Learners and Reluctant Users

Every teacher, especially at secondary level, have met at least one “reluctant learner” in his or her career. Reluctant learners are those students who achieve low academic results as a consequence of their scarce motivation, self-esteem and low efficacy. They are usually disengaged with school and they do not easily get involved in learning activities; they often end up leaving school without any formal qualification. Since Mathematics is considered one of the hardest disciplines, developing negative aptitudes towards it is rather more common than other subjects [25]. The use of digital learning environments, in particular for formative assessment, is generally considered successful to engage reluctant learners in Mathematics. In fact, the presence of digital and dynamic elements can offer them new ways of engaging with mathematical thinking and of facilitating understanding [26].

In this framework, we propose a distinction between the reluctant *learners*, who are dissatisfied with school, or with Mathematics, in general, and reluctant *users*, who cannot be engaged in the use of learning technologies. In particular, we define “reluctant learners” as those students who make too few attempts to the AFA activities in a course to benefit of their effects on learning, metacognition and self-regulation. They are counterposed to the “active users”, who can be actively involved in the online activities. Activity and reluctance here concern only the use of the automatically graded assignments; an a priori relation with learning is not assumed, however, it will be discussed through the following analysis.

3 Research Questions

The Department of Mathematics of the University of Turin (Italy) has been active for years in the research on automatic formative assessment. In light of the theoretical framework illustrated above, the Department has started a research project aimed to study the effectiveness of the learning approach based on the model of AFA and interactive feedback previously developed [22] in Mathematics education at lower secondary level. Our first goal was the study of the effectiveness of the AFA model on the point of view of learning, in order to understand the advantages of using these kinds of online activities for improving learning results. Interesting insights can be drawn from the comparison of an experimental group of classes which uses the AFA with a control group of classes which received traditional instruction, but also from the comparison between active users’ and reluctant users’ results. Secondly, we were interested in understanding to what extent the reluctant users of the online activities are reluctant learners, and what the reasons are that motivate their scarce use of the AFA.

These purposes led us to the formulation of the following research questions:

1. Are mathematical knowledge, procedural understanding and self-assessment skills influenced by the regular adoption of the formative automatic assessment for learning Mathematics at secondary school level?
2. Who are the reluctant users of the AFA and are there any differences between active and reluctant users in the learning achievements?

In the next paragraphs these research questions will be addressed and discussed through the results of a didactic experimentation.

4 Experimentation and Research Method

4.1 The Experimentation “Educating City”

The didactic experimentation on which this paper focuses is part of a bigger project, called “Educating City” (“Città Educante” in Italian), funded by the National Research Council, with the aim of rethinking the learning processes through the application of the newest advances in educational technologies [27]. Within the project the University of Turin undertook an experiment during the school year 2017/2018, involving 24 classes of 8th grade of 6 different lower secondary schools, with a total amount of 546 students, with their teachers of Mathematics. 13 of these classes (299 students), randomly chosen, formed the test group and they carried on several didactic activities proposed by experts from the Department; the remaining 11 classes (247 students) constituted the control group for the comparison of results. The six schools were located in different areas of Turin; about half of the classes belonged to a low socioeconomic status with a high presence of immigrants’ children, while the other half of the sample belonged to a middle-high social class and wealthier families. Of each school, a similar number of classes was inserted in the control group and in the test group, so that the two groups were homogeneous under this point of view.

The classes in the test group had access to an online course with automatically assessed digital materials for revising and learning mathematical contents relevant for the competences to be acquired by the end of 8th grade. The materials were prepared by experts from the University of Turin, after agreeing on the topics with their Mathematics teachers; the online course was then shared with teachers, who were trained on the use of the platform. Teachers could use the interactive materials in class during their lessons and assign the automatically assessed questions as homework. Moreover, a particular relevant topic in Mathematics education for grade 8, that is formulas, symbolic manipulation and modeling, was chosen as didactic goal for the experimentation, and a series of lessons around it were held in collaboration with a PhD student of the Department.

4.2 Instruments

All the students of both test and control classes took a test at the beginning of the school year, an intermediate test halfway through and a final one in June. The initial test was aimed to gain a picture of students’ abilities and competence in Mathematics; it was composed of 19 items with different levels of difficulty and varied topics. The test was of an acceptable level of reliability (Cronbach’s alpha: 0.773). The intermediate test was mainly focused on symbolic manipulation and modeling, and it was administered before starting that particular module, so that it measured students’ prior modeling skills. It was composed of 25 items with different levels of difficulty, all on the same topic; its Cronbach’s alpha is 0.800, revealing a high level of internal

reliability. The final test aimed to assess high-level mathematical competences that students should develop during the lower secondary school; it was composed of 11 items, all of them rather difficult; its level of reliability is acceptable, considering that the number of items was lower (Cronbach's alpha: 0.652). During the final test students were also asked to fill a self-assessment form after solving the exercises: for each question, they were asked to indicate if they thought that their answer was correct, incorrect or whether they did not know it. All the tests were paper-and-pen based so that students in the test group would not be favored by their experience with the use of the AAS.

At the beginning of the experimentation all the students filled in a questionnaire about their motivation and approach to school in general and to Mathematics in particular; one more questionnaire was distributed only to the test classes at the end of the school year to investigate the appreciation of experimental methodologies. Questionnaires were delivered online that were mainly composed of Likert-scale questions, with a few open questions for comments and explanations.

4.3 The Automatically Assessed Learning Materials

The online course prepared for the experimentation was created in a Moodle platform dedicated to the project (<https://cittaeducante.i-learn.unito.it/>). The course was organized in 10 modules, corresponding to 10 mathematical topics; an additional initial section included introductory materials and interactive tutorials on how to use the platform and how to insert answers in the online tests. The course was replicated in 13 separated courses so that each class could constitute a virtual community, easily monitored by its teacher. In each section there were two kinds of materials: problems with interactive solutions that teachers could use in class with the IWB (Interactive White Board) and online assignments conceived for students' homework. The latter are the main focus of this paper, while classroom problem solving activities will not be discussed here. Teachers were free to use all the sections that they needed; materials about "formulas and functions", the most considerable section for the topic involved and for the quantity of materials, were compulsory for all the classes in the test group. Working with formulas and functions is one of the most important part of 8th grade Mathematics and it was deeply dealt in the control classes as well. Teachers of the control group were not provided with didactic materials; they were asked to teach their lessons in a traditional way covering all the topics as they usually do with 8th grade classes. Figure 1 shows the homepage of the online course with 10 topic sections and an initial one; Fig. 2 shows the content of one section, with problems and the related automatically graded assignments.

The online assignments were created through Moebius Assessment. The choice of this system is motivated by its powerful assessing capabilities for STEM, long known by the main scientific universities [17] and recently developed with new interactive functionalities and a simple interface that makes it usable even at school level; moreover, it has been successfully adopted by the University of Turin for online courses and in several projects involving secondary schools [28, 29], so that the research group involved in this experimentation has gained deep experience in the technical and methodological use of this AAS.



Fig. 1. Online course available to teachers and students of the test group.

Lettere e aree	← Problem
<input checked="" type="checkbox"/> Calcola l'area delle figure	
<input checked="" type="checkbox"/> Componi e scomponi le figure	
<input checked="" type="checkbox"/> Ancora formule e aree	← Online assignments
<input checked="" type="checkbox"/> Altre strane figure	
<input checked="" type="checkbox"/> Dalle aree ai grafici	
Lettere e volumi	
<input checked="" type="checkbox"/> Volumi, superfici e grafici	
Una lettura per le vacanze	
<input checked="" type="checkbox"/> Altre letture estive	
Esercizi per le vacanze	

Fig. 2. One section's content

The assignments were created according to the above-mentioned model of formative assessment. Questions involved a mathematical task, which was often contextualized in reality or in another discipline; verbal and symbolic registers were associated with graphical visualizations in two or three dimensions, tabular representations to be completed by students and other interactions that fostered understanding. The feedback was given interactively either after giving the incorrect answer to the main task, or through a series of subtasks that were necessary parts of the procedure that helped students check if their computations were correct before inserting the final answer. As an example, Fig. 3 shows a question where students were guided into drawing a line in the cartesian plane. As a first step they were asked to fill in a table with the coordinates of

✔ We are going to represent the line of equation $y = -2x + 1$ on the Cartesian plane.
 Let's start from finding **two** points belonging to the graphic of the line.
 Two points are sufficient to identify one line.

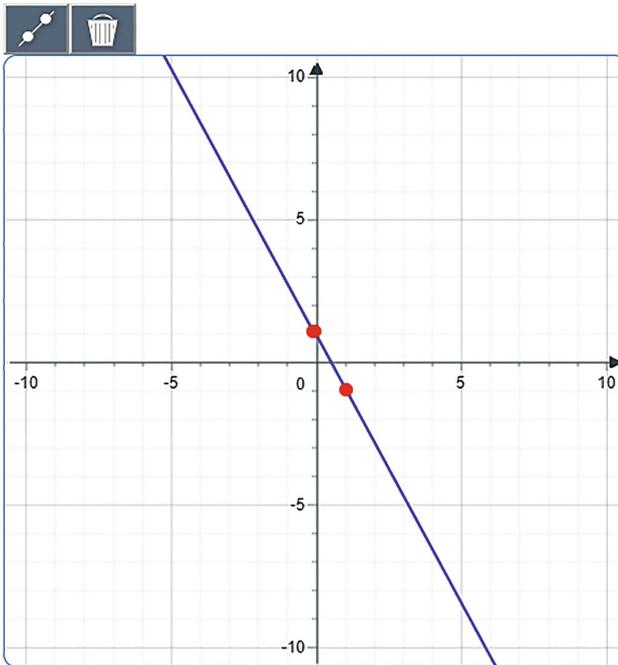
x	y
0	1 ✔
1	-1 ✔

Click on the **Verify** button to check your answer and draw the line.

Attempt 1 of 1

Verify

Now draw the line in the following Cartesian plane. To sketch the line, click on two points of the Cartesian plane.



Click on the **Verify** button to check your answer.

Attempt 1 of 1

Verify

Fig. 3. An example of question made through the AAS Maple TA. The Italian text has been translated into English for the paper's comprehension.

two points belonging to the graphic of the line; once they checked that the values were correct, they were asked to interactively draw the line in the cartesian plane. A feedback showing the correct line was then provided to all students after their sketch was graded. At the following attempt, another line with a different equation would be displayed. Each assignment was made up of no more than three questions of this kind, or even only one question if it involved a more complex problem.

4.4 Data Collection

In order to discuss the answers to the research questions, five main sources of data were collected and cross-checked:

1. the percentage of correct answers of the initial, intermediate and final tests returns the information about the learning achievements made by the two groups of students (test and control group);
2. answers from the self-assessment form expressing students' self-assessment skills. For each item, one point was assigned if the student's impression of giving the correct or incorrect answer matched the item's evaluation, while no points were assigned in case it did not match or in case students were unsure of their answers. The sum of the points earned in the whole test was then scaled from 0 to 10;
3. students' answers to the initial questionnaire indicate their motivations to study Mathematics – and to schoolwork in general – before the experimentation;
4. the answers of the test group to the final questionnaire – and in particular to the set of questions about the online assignments – shows students' satisfaction with the online interactive materials;
5. data about the platform usage have been extracted, in particular the number of attempts made and the grades obtained, to analyze the relationship between learning achievements and frequency of usage.

The last point deserves some additional clarifications. For each student, the number of attempts to all the available tests – included repeated attempts to the same tests – was drawn from the platform. Only the fully completed attempts were considered as “attempt”, meaning that students requested and obtained feedback and grade. This choice is in line with the value of the formative feedback proposed in this discussion, as we maintain that without automatic grading and interactive feedback, students would not appreciate all the potentialities offered by the system, and this is what makes this assessment different from a traditional one. Thus, all the times students opened a test just to have a look without completing it were not taken into account.

The number of attempts was used to define the two sub-groups of students in the test group: we defined “active users” as those who made at least 5 graded attempts for the assignments in the course, while the students who made less than 5 attempts were defined “reluctant users”. The latter represent about 25% of the students in the test group. As active and reluctant users were defined on the basis of the number of accesses to online materials, a parallel definition is not provided for students belonging to the control group, who did not have access to online activities. Data have been analyzed using SPSS 25; the main results and conclusions will be reported in the following paragraphs.

5 Results

The main results of the experimentation are positive, as the students in the test group performed significantly better in the final test than their school-mates in the control group. The mean of the final test results, corrected with the initial test results through covariance analysis, are shown in Table 1. Values are statistically significant ($p = 0.002$). The sample has been restricted to the only students who took both the initial and the final test, being present at school on the days the tests were administered.

Table 1. Corrected means of the final test's results.

	N	Mean	Standard error
Control group	197	43.99	1.185
Test group	255	49.01	1.040

In order to better understand the relationship between active users and reluctant users in the test group with regard to learning gains, the analysis of variance (ANOVA) has been performed on their results in the initial, intermediate and final tests and in the self-assessment of the final test. Results are summarized in Table 2; they include 62 reluctant users and 193 active ones. It emerges that active users performed a little, but not significantly better in the initial and intermediate tests than reluctant users. However, active users performed significantly better in the final test and in the self-assessment of the final test. Given the similarity in the initial results, it is not possible to reach the conclusion that reluctant users were initially weaker than their active classmates; the regular use of the automatic formative assessment, however, led to higher learning achievements and to more effective self-assessment skills.

Table 2. Comparison between learning results of reluctant and active users.

		Reluctant users	Active users	p-value
Initial test	Average	36.94	42.51	0.060
	St. dev.	19.26	19.12	
Intermediate test	Average	55.41	59.85	0.091
	St. dev.	17.07	16.37	
Final test	Average	43.16	49.13	0.049
	St. dev.	18.20	20.21	
Self-assessment	Average	3.64	4.47	0.011
	St. dev.	2.09	2.13	

An analysis on the platform usage leads to affirm that the questions implemented with automatic assessment according to the above-mentioned educational models were effective under several points of view. The number of attempts that students made to any test is, on average, 15.81; this value obviously grows when considering only active

users (average: 20.76). Active students attempted the same assignments an average rate of 1.75 times, which means that they have generally made more than one attempt to the online quizzes. The feedback gained with the automatic assessment has been used to improve the performance in subsequent attempts, where tasks were repeated with different numerical values, functions, and geometrical figures which varied algorithmically. This can be statistically shown using the assignments results from the platform, comparing the average score of students' first assignment attempts with the average score of their last attempt. Active students' results grew from an average value of 51.97 at their first attempts to 60.25 at their last attempts (values are expressed in percentage of correct answer); a pairwise t-test shows that the difference is statistically significant ($p < 0.001$). It is interesting to notice that the average grade of the first attempts made by reluctant users (49.76), is similar to that of active students: the ANOVA test shows that there is no statistically significant difference between the two means ($p = 0.524$). This fact confirms what emerged from the initial and intermediate tests, namely that reluctant users were not initially weaker than active users in terms of learning. However, in the platform, like in the learning tests, reluctant users did not improve their results: the average of their last attempts increased only of 2.15 points out of 100, rising up to 51.91, while the average active users' grade increased by 8.28 points out of 100, rising up to 60.25. The difference in the improvements between the two groups can be seen as a consequence of the time spent in the use of the online assignments: reluctant users' average rate of attempts per assignment is 1.22. Table 3 summarizes the statistics on the platform usage for active and reluctant users.

Table 3. Comparison of data from platform usage between active and reluctant users.

		Reluctant users	Active users	p-value
Number of attempts	Average	1.48	20.76	<0.001
	St. dev.	1.52	15.63	
Rate of attempts per test	Average	1.22	1.75	0.001
	St. dev.	0.43	0.90	
Average of the first attempts	Average	49.76	51.97	0.524
	St. dev.	24.90	18.49	
Average of the last attempts	Average	51.91	60.25	0.016
	St. dev.	22.82	18.80	

The improvements registered through the platform suggest that the automatic assessment format helped students develop problem solving strategies, acquire procedural understanding, identify and correct their mistakes. The answers of active users to the final questionnaire support this hypothesis: they affirmed that it was useful to visualize the feedback immediately after giving the answer, and in particular it helped them understand how to solve a task; step-by-step guided resolutions were useful to understand the solving process. Their answers are compared with the reluctant users' ones, which are lower, as expected. Interestingly, ANOVA results show that, between the two groups, there is no statistically significant difference neither in the appreciation

of AFA as a support to study, nor in the effectiveness of AFA to raise the awareness of one's capabilities, while the main differences in appreciation are due to the ability of the system to help learn a method to solve problems. It means that reluctant users acknowledged the potential of the automatic assessment; anyway, the effectiveness of the system to learn problem solving strategies could actually be appreciated only after multiple attempts and after an adequate time spent on online homework. The data are summarized in Table 4.

Table 4. Active users' answers to questions about automatic assessment in the final questionnaire. Answers are in a Likert scale from 1 (lowest value) to 5 (highest value).

		Active users	Reluctant users	p-value
It is useful to visualize the correct answer immediately after giving a response	Av.	4.29	3.93	0.047
	S.D.	0.89	0.91	
The immediate assessment helped me understand how to answer the questions	Av.	4.16	3.77	0.042
	S.D.	0.95	0.97	
Problems with step-by-step guided resolution helped me understand how to solve the exercises	Av.	3.91	3.47	0.025
	S.D.	0.96	1.01	
Online assignments helped me learn to solve problems autonomously	Av.	3.34	2.83	0.015
	S.D.	1.00	1.08	
Online tests are a valid support for studying	Av.	3.62	3.43	0.362
	S.D.	0.99	1.07	
Online assignments helped me acquire awareness of my preparation	Av.	3.69	3.43	0.20
	S.D.	0.99	1.00	

Students' appreciation of the automatic assessment is confirmed by their answers to the open question "Why do you think that online tests are useful for learning Mathematics?". Examples of comments repeated in the answers collected are the following:

- "I appreciated the fact that the exercises were on the computer, which made them more interesting. In my opinion, giving an evaluation to each exercise pushed me to put more effort in my homework and to improve my results."
- "They are useful because if you give an incorrect answer you can try it again, while you cannot do the same on the book".
- "The online exercises are useful because you can access them whenever you need. Also, the problems that gave the correct solution immediately after your answer were excellent, since you could understand the solving process and try them again with different data."
- "In my opinion, they are useful because they help you understand the topic, and the presence of figures and graphs helped you better understand the task."

As using the automatic assessment provides such good effects on learning, it seems a pity that there were students who missed this opportunity and barely opened the online assignments. Data collected during the experimentation help understand who the reluctant users are. As emerged above, reluctant users do not coincide with reluctant

learners, in fact their initial learning achievements were not significantly far from the active users' ones. We supposed that students who made little use of the online platform had a preference for a paper-and-pen learning approach, so we analyzed their answers to the question inserted in the final questionnaire: "do you prefer doing computer-based or paper-based homework?", to which students could choose one of the two options. It turned out that the 60% of active users opted for "computer-based homework", while 45% of reluctant users made the same choice (Cramer's V: 0.117, p-value: 0.11). This means that, among reluctant users, there is a slight preference for a paper-and-pen learning style, but it does not explain the difference in the platform's usage. At least, the 45% of reluctant users who stated that they preferred using the computer for doing their homework, actually did not do it.

We investigated reluctant users' general approach to learning through their answers to the initial questionnaire, provided in a Likert scale from 1 (strongly disagree) to 4 (strongly agree) and displayed in Table 5. It emerges that they are as aware as active users of the importance of studying Mathematics for their educational and professional future; however, they do not appreciate the subject per se as much as their classmates (Cramer's V: 0.22, $p = 0.006$). Some differences emerge in the extent to which they appreciate the way Mathematics is taught at school: reluctant users are slightly less willing to attend Mathematics lessons. They are aware that good results depend on their dedication to the same extent of their classmates, they arrive at school with their homework done, however they work less hard on their homework and they usually do not help their friends with homework as much as active students (Cramer's V: 0.24, $p = 0.002$). There are no statistically significant differences in questions asking about their attitude towards problem solving, even though reluctant users seem to be a little less persistent in problem solving.

Table 5. Answers of reluctant and active users to questions of the initial questionnaire, given in a Likert scale from 1 (strongly disagree) to 4 (strongly agree).

		Reluctant users	Active users	Cramer's V (p-value)
Mathematics is an important subject as it will be useful for my future studies	Av.	3.43	3.46	0.15 (0.093)
	S.D.	0.61	0.75	
It is worth working hard in Mathematics because it will be useful for my professional future	Av.	3.46	3.51	0.14 (0.167)
	S.D.	0.53	0.66	
I'm looking forward to having Mathematics lessons	Av.	2.20	2.45	0.151 (0.116)
	S.D.	0.73	0.81	
I study Mathematics because I like it	Av.	2.58	2.74	0.22 (0.006)
	S.D.	0.83	0.88	
If I work hard, I can be successful in Mathematics	Av.	3.42	3.42	0.086 (0.595)
	S.D.	0.68	0.66	
I finish all my Mathematics homework before the lessons	Av.	2.94	2.94	0.073 (0.716)
	S.D.	0.97	1.03	

(continued)

Table 5. (continued)

		Reluctant users	Active users	Cramer's V (p-value)
I work hard on my Mathematics homework	Av.	2.69	2.94	0.148 (0.129)
	S.D.	0.79	0.69	
I help my friends in Mathematics	Av.	1.89	2.29	0.244 (0.002)
	S.D.	0.89	0.86	
I get easily discourage when I try to solve a problem	Av.	2.52	2.32	0.107 (0.565)
	S.D.	1.17	1.14	
I put aside difficult problems	Av.	2.45	2.20	0.145 (0.246)
	S.D.	1.25	1.04	

Sociocultural factors did not affect students' usage of the automatic assessment: the percentage of reluctant users is similar and even lower in schools located in areas with low socioeconomic conditions (24%) than in wealthy areas (27.1%).

Summing it up, the picture of the reluctant users emerging from data is that of students who are not lacking talents or economic conditions for accessing the platform or performing well; they are conscious that school is important even though they do not like Mathematics. This lack of inner motivation could be linked to the scarce effort they put in studying. An important element of distinction of reluctant users from the general reluctant learners is their acknowledgement that school success depends on their effort in studying. It is interesting that they hardly help their friends with Mathematics: associated to the fact that they complete their homework without working hard on it, it suggests that they are mainly the students who are usually helped by others or even copy their homework, a thing that becomes nearly impossible with an online system. This insight was also confirmed by teachers' opinions collected through a focus group at the end of the experimentation: they affirmed that the online work was really appreciated by the weakest students with learning disorders or little self-awareness; it was useful for skilled students to improve even more, however the most undisciplined and lazy students did not seize the opportunities given by digital tools, they worked very little on the platform, just like they did with paper-and-pen homework.

6 Main Conclusions and Implications for Research and Practice

This study has several limitations that hinder the generalization of the results: numbers are significative, but not that high; moreover, all the participants come from the same city in Italy: regional differences are not taken into account in this study. However, it offers some interesting insights that deserve to be deepened in future research.

The data collected in the experimentation help answer the research questions: (1) Are mathematical knowledge, procedural understanding and self-assessment skills influenced by the regular adoption of the formative automatic assessment for learning Mathematics at secondary school level? and (2) Who are the reluctant users of the AFA and are there any differences between active and reluctant users in the learning achievements?

Regarding the first research question, it resulted that the immediate feedback, given at the right time and in an interactive way, helps students master a process for solving mathematical tasks and find problem solving strategies. In fact, when students repeated the exercises with similar tasks but different data, they improved their grades, as shown in the analysis before (Table 3). This finding is rather in disagreement with many findings in literature, which show that students pay little attention to feedback information and consequently do not use it to improve [30]; in this study, the key strength was the interactivity of the feedback, which managed to engage students in processing the information. Enhancements in understanding also emerged from the comparison of initial and final tests, compared with other students in similar conditions who did not use the platform. The gains involve not only mathematical knowledge, but also metacognitive factors, self-assessment skills above all. This fact emerged clearly from the performance in the self-evaluation of the final test of students that regularly used the automatic assessment, when compared with that of students who rarely worked on the platform. Hence, we have positively answered the first research question: mathematical knowledge, procedural understanding and self-assessment skills are positively influenced by the regular adoption of the formative automatic assessment for learning Mathematics at secondary school level.

Regarding the second research question, through an analysis of the data usage, we have selected a group of students who made few or no attempts with the online assignments and we defined them “reluctant users”, opposed to the “active users”. Data show that students that did not take advantage of this chance did not improve in learning as much as their classmates. Their profile can be inferred from their answers to the initial questionnaire, thus answering the second research questions. Reluctant users are basically students who put little effort in school in general, and who dislike Mathematics in particular. Economic disadvantages or learning disorders did not prevent students from working in the online courses, as reluctant users’ social distribution and their initial results show. Some of them prefer a paper-based approach to learning, but it is not the main reason why they did not use the technologies.

The results of this experimentation lead to consider other aspects of the integration of the automatic assessment in Mathematics courses at school level. First of all, teachers need to be prepared to carry out this step. The materials included in the online course presented in this paper were all designed and developed by university researchers, who are expert both in the technical use of the platform and in didactic methodologies for teaching and learning Mathematics. The teachers who participated to the experimentation, satisfied with the impact of this kind of technology in their classes, started a training aimed at making them autonomous in authoring digital materials of this kind. For a similar training to be effective, teachers should have constant support by expert tutors, and inner motivation to change their teaching method. All the online activities, which resulted to be successful in delivering learning, have been made available to all the Italian teachers through the national project “Problem Posing and Solving”, supported by the Italian Ministry of Education, where didactic methodologies and technologies as the automatic formative assessment are proposed and used by teachers supported by tutors [21]. Moreover, the results of this experimentation gave prompts to the research on adaptive assessment about which the department of

Mathematics of the University of Turin is very active [31, 32]. The main output of this research is the design of an automatic system that provides students with questions according to their level of competence, choosing them among a database of shared items appositely clustered. Our hope is to broaden the set of good experiences and examples of learning through technologies by means of the cooperation of research, training and practice, in order to offer more and more students effective opportunities of education.

References

1. World Economic Forum: Eight Futures of Work. Scenarios and their Implications (2018)
2. Drijvers, P., Ball, L., Barzel, B., Heid, M.K., Cao, Y., Maschietto, M.: Uses of Digital Technology in Lower Secondary Mathematics Education: A Concise Topical Survey. Springer, New York (2016). <https://doi.org/10.1007/978-3-319-33666-4>
3. Olive, J., Makar, K.: Mathematical Knowledge and Practices Resulting from Access to Digital Technologies. In: Hoyles, C., Lagrange, J.-B. (eds.) Mathematics Education and Technology-Rethinking the Terrain. NISS, vol. 13, pp. 133–178. Springer, New York (2010). https://doi.org/10.1007/978-1-4419-0146-0_8
4. Heid, M.K.: The technological revolution and the reform of school mathematics. *Am. J. Educ.* **106**(1), 5–61 (1997). <https://doi.org/10.1086/444175>
5. Stacey, K., Wiliam, D.: Technology and assessment in mathematics. In: Clements, M., Bishop, A., Keitel, C., Kilpatrick, J., Leung, F. (eds.) Third International Handbook of Mathematics Education. SIHE, vol. 27, pp. 721–751. Springer, New York (2013). https://doi.org/10.1007/978-1-4614-4684-2_23
6. ECDL Foundation: The Fallacy of the ‘Digital Native’: Why Young People Need to Develop their Digital Skills (2014)
7. Mitra, A., Steffensmeier, T.: Changes in student attitudes and student computer use in a computer-enriched environment. *J. Res. Comput. Educ.* **32**(3), 417–433 (2000)
8. Armstrong, D.A.: Students’ perceptions of online learning and instructional tools: a qualitative study of undergraduate students use of online tools. *Turk. Online J. Educ. Technol.* **10**(3), 222–226 (2011)
9. Beevers, C.E., Cherry, B.S.G., Clark, D.E.R., Foster, M.G., McGuire, G.R., Renshaw, J.H.: Software tools for computer-aided learning in mathematics. *Int. J. Math. Educ. Sci. Technol.* **20**(4), 561–569 (1989)
10. Sangwin, C.: Computer aided assessment of mathematics using STACK. In: Cho, S.J. (ed.) Selected Regular Lectures from the 12th International Congress on Mathematical Education, pp. 695–713. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-17187-6_39
11. Barana, A., Marchisio, M., Rabellino, S.: Automated assessment in mathematics. In: Proceedings of 2015 IEEE 39th Annual Computer Software and Applications Conference, pp. 670–671. IEEE, Taichung (2015). <https://doi.org/10.1109/COMPSAC.2015.105>
12. Black, P., Wiliam, D.: Developing the theory of formative assessment. *Educ. Assess. Eval. Account.* **21**(1), 5–31 (2009)
13. Hattie, J., Timperley, H.: The power of feedback. *Rev. Educ. Res.* **77**(1), 81–112 (2007)
14. Sangwin, C., Makar, K., Cazes, C., Lee, A., Wong, K.L.: Micro-level automatic assessment supported by digital technologies. In: Hoyles, C., Lagrange, J.-B. (eds.) Mathematics Education and Technology-Rethinking the Terrain. NISS, vol. 13, pp. 227–250. Springer, New York (2010). https://doi.org/10.1007/978-1-4419-0146-0_10

15. Paiva, R.C., Ferreira, M.S., Mendes, A.G., Eusébio, A.M.J.: Interactive and multimedia contents associated with a system for computer-aided assessment. *J. Educ. Comput. Res.* **52**(2), 224–256 (2015)
16. Beevers, C.E., Paterson, J.S.: Automatic assessment of problem solving skills in mathematics. *Act. Learn. High Educ.* **4**(2), 1–14 (2003)
17. Bokhove, C., Drijvers, P.: Digital tools for algebra education: criteria and evaluation. *Int. J. Comput. Math. Learn.* **15**(1), 45–62 (2010)
18. Helder, E., Sosnovsky, S., Dimitrova, V.: Adaptive intelligent learning environments. In: Duval, E., Sharples, M., Sutherland, R. (eds.) *Technology Enhanced Learning: Research Themes*, pp. 109–114. Springer, New York (2017). https://doi.org/10.1007/978-3-319-02600-8_10
19. Nicol, D., Milligan, C.: Rethinking technology-supported assessment practices in relation to the seven principles of good feedback practice. In: *Innovative Assessment in Higher Education*, pp. 1–14. Taylor and Francis Group Ltd., London (2006)
20. Barana, A., Fioravera, M., Marchisio, M., Rabellino, S.: Adaptive teaching supported by ICTs to reduce the school failure in the project “Scuola Dei Compiti”. In: *Proceedings of 2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC)*, pp. 432–437. IEEE (2017). <https://doi.org/10.1109/COMPSAC.2017.44>
21. Brancaccio, A., Marchisio, M., Palumbo, C., Pardini, C., Patrucco, A., Zich, R.: Problem posing and solving: strategic italian key action to enhance teaching and learning mathematics and informatics in the high school. In: *Proceedings of 2015 IEEE 39th Annual Computer Software and Applications Conference*, pp. 845–850. IEEE (2015)
22. Barana, A., Conte, A., Fioravera, M., Marchisio, M., Rabellino, S.: A model of formative automatic assessment and interactive feedback for STEM. In: *Proceedings of 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)*, pp. 1016–1025. IEEE (2018). <https://doi.org/10.1109/COMPSAC.2018.00178>
23. Sadler, D.R.: Formative assessment and the design of instructional systems. *Instr. Sci.* **18**(2), 119–144 (1989)
24. Beevers, C.E., Wild, D.G., McGuine, G.R., Fiddes, D.J., Youngson, M.A.: Issues of partial credit in mathematical assessment by computer. *Res. Learn. Technol.* **7**(1), 26–32 (1999)
25. Sanacore, J.: Turning reluctant learners into inspired learners. *Clear. House* **82**(1), 40–44 (2008)
26. Calder, N., Campbell, A.: Using mathematical apps with reluctant learners. *Digit. Exp. Math. Educ.* **2**(1), 50–69 (2016)
27. Barana, A., Boffo, S., Gagliardi, F., Marchisio, M.: *Problem Posing & Solving: a Digital Way to Learn Mathematics* (in press)
28. Barana, A., Bogino, A., Fioravera, M., Marchisio, M., Rabellino, S.: Open platform of self-paced MOOCs for the continual improvement of academic guidance and knowledge strengthening in tertiary education. *J. E-Learn. Knowl. Soc.* **13**(3), 109–119 (2017). <https://doi.org/10.20368/1971-8829/1383>
29. Barana, A., Marchisio, M., Miori, R.: MATE-BOOSTER: design of an e-learning course to boost mathematical competence. In: *Proceedings of the 11th International Conference on Computer Supported Education (CSEDU 2019)*, pp. 280–291 (2019)
30. Timmers, C., Veldkamp, B.: Attention paid to feedback provided by a computer-based assessment for learning on information literacy. *Comput. Educ.* **56**(3), 923–930 (2011)

31. Marchisio, M., Di Caro, L., Fioravera, M., Rabellino, S.: Towards adaptive systems for automatic formative assessment in virtual learning communities. In: Proceedings of 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), pp. 1000–1005. IEEE (2018). <https://doi.org/10.1109/COMPSAC.2018.00176>
32. Barana, A., Di Caro, L., Fioravera, M., Marchisio, M., Rabellino, S.: Ontology development for competence assessment in virtual communities of practice. In: Penstein Rosé, C., et al. (eds.) AIED 2018. LNCS (LNAI), vol. 10948, pp. 94–98. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-93846-2_18