

Article

Interactive Feedback for Learning Mathematics in a Digital Learning Environment

Alice Barana *, Marina Marchisio * and Matteo Sacchet *Department of Molecular Biotechnology and Health Sciences, University of Turin, Via Nizza 52,
10126 Torino, Italy* Correspondence: alice.barana@unito.it (A.B.); marina.marchisio@unito.it (M.M.); matteo.sacchet@unito.it (M.S.)

Abstract: The COVID-19 pandemic has evidenced a need for tools and methodologies to support students' autonomous learning and the formative assessment practices in distance education contexts, especially for students from challenging backgrounds. This paper proposes a conceptualization of Interactive Feedback (IF) for Mathematics, which is a step-by-step interactive process that guides the learner in the resolution of a task after one or more autonomous tentative. This conceptualization is grounded on theories and models of automatic assessment, formative assessment, and feedback. We discuss the effectiveness of the IF for engaging students from low socio-economic contexts in closing the gap between current and reference performance through a didactic experimentation involving 299 Italian students in grade 8. Using quantitative analyses on data from the automatic assessment, we compared the results of the first and last attempts in activities with and without IF, based on algorithmic parameters so that the task changes at every attempt. We found that IF was more effective than other kinds of activities to engage learners in actions aimed at improving their results, and the effects are stronger in low socio-economic contexts.

Keywords: automatic assessment; digital learning environment; formative assessment; interactive feedback; learning mathematics; mathematics education



Citation: Barana, A.; Marchisio, M.; Sacchet, M. Interactive Feedback for Learning Mathematics in a Digital Learning Environment. *Educ. Sci.* **2021**, *11*, 279. <https://doi.org/10.3390/educsci11060279>

Academic Editors: Giorgio Bolondi, Federica Ferretti and Luis J. Rodríguez-Muñiz

Received: 27 April 2021
Accepted: 3 June 2021
Published: 5 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Digital technologies can offer powerful tools and environments to enhance the practice of formative assessment in Mathematics Education [1,2]. The use of digital learning environments for assessment enables the generation and collection of data about learning agents (such as teachers, students, tutors, peers), processes, and results. These data can be used to drive and adjust the learning path, make choices and decisions, and support learning in several ways [3,4].

The last year has been marked by the consequences of the COVID-19 health emergence, which upset and permanently transformed the educational processes, requiring a sudden and urgent revision of the paradigms of the educational processes. In particular, the school systems evidenced a need for innovative solutions to support the transition of assessment and formative assessment practices for home-schooling and distance education [5–7]. Several recent studies are highlighting that Mathematics is the subject in which the learning loss due to the pandemic is the most evident; the decrement in learning seems to be unevenly distributed, with students from less privileged backgrounds hit harder than others [8,9]. These data are alarming since Mathematical competence is relevant for the development of lifelong learning and active citizenship [10]. Especially young students learning from their home feel the need of a guide in their school-related activities. They need tools to drive their cognitive processes and develop competences also without the physical presence of a teacher. In this situation, online formative assessment can be helpful to develop self-regulation competence, a relevant skill related to learning how to learn and related to life-long learning, and essential in distance learning to make students

autonomous in driving their learning and problem-solving processes [11,12]. Interactive learning environments, where the student's action is encouraged and where what happens next depends on this action [13], should be promoted and integrated into the teaching practices, both in classroom-based and online situations, not to create discontinuities in the adopted methodologies and instruments depending on the situation [14].

In this paper, we propose a model of interactive feedback for Automatic Formative Assessment of Mathematics. We conceptualize Automatic Formative Assessment (AFA) as the practice of formative assessment in a Digital Learning Environment through the automatic elaboration of students' answers and provision of feedback. AFA can be used in a Digital Learning Environment in classroom-based, blended, hybrid, or online activities [15]. We conceptualize Interactive Feedback (IF) as a step-by-step interactive process guiding the learner in the resolution of a task after one or more autonomous attempts. Our model is supported by theories on formative assessment and feedback; it requires the use of an automatic assessment system suitable to assess mathematical objects. Even though the literature offers several examples of step-by-step activities for problem-solving in Mathematics [16–18], our innovative contribution is conceptualizing the whole interactive solving process as a feedback and embedding it in the framework of Automatic Formative Assessment. Through a didactic experimentation involving 299 students in grade 8, we have analyzed the effectiveness of the interactive feedback to improve learning results in Mathematics.

This paper is structured as follows: Section 2 (theoretical framework) discusses the background theories and our conceptualization of interactive feedback, with some examples; Section 3 (research question and method) presents the research question and the design of the experimentation, giving details on the analyses performed on data from automatic formative assessment activities; in Section 4 (results) the results of the analyses are shown with reference to the theoretical framework in order to answer the research questions; in Section 5 (discussion and conclusions) the results of the analyses are commented on, suggesting the potential impact of the adoption of the model of interactive feedback at a largest scale.

2. Theoretical Framework

In this section, we outline the main points of the theories on which we frame our research, namely: formative assessment and feedback. Then, we provide a definition and a characterization of automatic formative assessment and interactive feedback.

2.1. Formative Assessment

In the last decades, the concern about formative assessment has grown to cover one of the major educational research issues. The contributions of Paul Black and Dylan Wiliam stood out in the development of a theoretical framework for formative assessment. We accept the definition they gave in [19]:

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.

This definition entails the collection of evidence, which can be gathered through tasks or questions, and the interpretation and use of the pieces of information gathered to act on learning. According to this definition, the mere collection of students' answers, without using them to make decisions to tailor their learning path, is not to be considered as formative assessment. Wiliam clarified this concept in another paper in 2006 [20], where he stated that "assessments are formative [. . .] if and only if something is contingent on their outcome, and the information is actually used to alter what would have happened in the absence of the information." To use the information gathered from assessment during teaching, it is essential to create moments of contingency. They are points in the instructional sequence where the instruction can change direction in light of evidence about the students'

achievement. This allows teachers to adapt the educational path to meet better students' learning needs [20]. These moments can be synchronous, including teachers' real-time adjustments during a classroom discussion after students' answers, or asynchronous, such as the use of evidence from students' homework to plan the next lesson. It seems that asynchronous moments of contingency are less effective, maybe due to the scarce experience of teachers with formative practices, or because of the use of inadequate tasks, or even due to the teachers' pressure to complete the program, to the disadvantage of students' understanding [20].

Black and Wiliam [19] identified three agents that are principally activated during formative practices: the teacher, the student, and peers. All three agents can be the subject of the decision-making process which follows the collection of evidence and which is at the core of formative assessment. Depending on the learning environment and conditions, the three can be more or less involved in the assessment process stages. Black and Wiliam [16] further developed a framework of formative assessment, individuating three different processes of instruction, that are the following:

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

Moreover, the two researchers theorized five key strategies enacted by the three agents during the three different processes of instruction:

1. Clarifying and sharing learning intentions and criteria for success;
2. Engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding;
3. Providing feedback that moves learners forward;
4. Activating students as instructional resources; and
5. Activating students as the owners of their own learning.

2.2. Feedback

The provision of feedback is only one of many strategies for formative assessment; nonetheless, it is probably the most distinctive and object of in-depth studies. The power of feedback emerges in Hattie's metanalysis [21]: with an effect size of 0.73, it results in one of the most effective strategies for learning. However, in the literature, results on feedback efficacy on learning are controversial [22]; for instance, one of the most surprising results that emerged from Kluger and DeNisi's review on feedback [23] is that in more than one-third of the 607 analyzed cases (effect sizes), feedback interventions reduced performance. This means that much attention should be paid to the feedback and the task's design. In this paper, we use Hattie and Timperley's definition of feedback: "*information provided by an agent (e.g., teacher, peer, book, parent, self, experience) regarding aspects of one's performance or understanding*" [24]. In that context, they provided a model for constructing effective feedback. The purpose of feedback is to reduce the discrepancy between current and desired understanding, and it can be fulfilled both by students and by teachers. Effective feedback must answer three main questions: "*Where am I going?*", "*How am I going?*", "*Where to next?*". In other words, they should indicate what the learning goals are (Feed Up); what progress is being made toward the goal (Feed Back); and what activities need to be undertaken to make better progress (Feed Forward). The three questions correspond to the three processes identified by Black and Wiliam in the model of formative assessment (establishing where the learners are in their learning, where they are going, and what needs to be done to get them there).

Feedback can work at four levels:

- The task level, giving information about how well the task has been accomplished;
- The process level, adding details about the main process needed to perform the task;
- The self-regulation level, activating metacognitive processes; and
- The self level, adding personal evaluations and affects about the learner.

A major concern raised by many authors is that learners often do not go through feedback. It is clear that if the learners do not process feedback, they lose all their potentialities [12,25]. Sadler introduced the idea that feedback only works when it is used to alter the gap between current and reference performance [26]. If the information is not or cannot be processed by the learner to produce improvements, it will not affect learning. We adopt Sadler's model of feedback, according to which, for feedback to be effective, students have to:

- a. Possess a concept of the standard (or goal, or reference level) being aimed for;
- b. Compare the actual (or current) level of performance with the standard; and
- c. Engage in an appropriate action, which leads to some closure of the gap.

This model is in line with the previously mentioned Black and Wiliam's theory of formative assessment.

Black and Wiliam argued that positive words of appreciation that concern the self level can encourage the learner to process the whole feedback and use the information gained [19]. Since feedback at the self level, such as praises and rewards, is shown to have a minimal effect [24], or even negative [23], on learning, this is the reason why they are included in a framework of good practices; however, we think that they should be avoided, and we will propose an alternative solution to this problem in the following paragraphs.

There has been a discussion in the literature aimed at comparing elaborated feedback and corrective feedback—those that just say if the answer is correct or not. Many studies show that the former are more useful to improve [25,27,28]. Elaborated feedback can refer to explanations of the correct solution, links to further reading materials, cues, suggestions, or their combinations [28]. In Mathematics, elaborated feedback is often in the form of a worked example, that is, a proposal of a resolution of the task or of a similar one [29]. The great part of elaborated feedback models that the literature proposes is static: students have to read them carefully and compare them with their results. Some studies also show that, more often than expected, students do not read them at all, especially if they perceive the task as too complicated or they do not receive the feedback in a timely manner [25]. To overcome this difficulty, Shute proposes to split feedback into manageable units, thus avoiding cognitive overload [28].

In practice, how should good feedback be? The first feature that emerges from the literature is that feedback on the mere correctness of an answer, which only says yes or no, acting only at the task level, is not much effective: instead, it should provide more details about the correct solution [28,30]. However, a simple static explanation of the correct solution belongs to a transmissive educational style. It does not meet the interactive learning environment features, as it does not actively engage the student. "*A good feedback causes thinking,*" Black and Wiliam affirm in their paper, "*In praise of educational research: formative assessment*" [31]. In an interactive learning environment, feedback should activate students who have to do something with it, elaborate it, and discuss it to link it to prior knowledge [32]. In this perspective, feedback should be an interactive process, a formative interaction that involves the learner and the feedback's agent (teacher, peers, or students themselves) and which influences cognition [19]. Such interactive feedback can work at the process level helping students to understand how the task should be solved. Moreover, during this interactive process, the external information provided through the feedback should help students generate inner feedback, which can be used to understand and fill the gap between current and desired performance [32]. In other words, external feedback should be transformed into internal feedback to affect subsequent learning. Internal feedback is a key element of self-regulation: in this sense, interactive feedback can activate self-regulation.

2.3. Automatic Formative Assessment and Mathematics

The definition of formative assessment that we have mentioned before can be adapted to consider the technologies' contribution. Among the many forms of assessment which can successfully take advantage of technologies (e.g., e-portfolios, self-assessment, peer

assessment, etc.) [33–35], we mainly focus on automatic assessment, where the technology is used to analyze the students' answers and to return feedback. According to Kluger and DeNisi [23], computerized feedback is more effective than human-delivered ones: this is one advantage of practicing formative assessment in e-learning.

We define Automatic Formative Assessment as the use of formative assessment in a Digital Learning Environment (DLE) through the automatic elaboration of students' answers and provision of feedback, where formative assessment is intended as in the Black and Wiliam's definition.

In Mathematics, AFA is widely used in DLEs to engage and motivate learners. In order to go beyond the multiple-choice modality, research centers and universities have developed systems that can process open-ended answers, expressed in different registers, from a mathematical perspective and establish if they are equivalent to the correct solutions [36–38]. In this research, we adopted Möbius Assessment (previously known as Maple T.A.). It relies on Maple Advanced Computing Environment's engine, which allows for computations, elaboration, advanced grading, and algorithms. Among all the existing automatic assessment systems, Möbius Assessment was selected for its high suitability for Mathematics, its powerful capabilities for assessment, and the possibility of integration within the most common Learning Management Systems such as Moodle. In Möbius Assessment, open mathematical answers can be implemented and graded through algorithms that verify if the student's answer matches the correct one independently of the form. This allows us to test different and complex cognitive processes, without being bound to the multiple-choice modality. Moreover, it is possible to create variables based on algorithms, random parameters and mathematical formulas, graphics, and even animated plots. Thus, questions appear different from student to student, and different at every attempt; data and solutions can be automatically computed in the algorithm within the question itself. In addition, the system supports adaptive capabilities, so that the next question part is proposed to the student according to the previous answer.

Several studies in the literature show how the adoption of Möbius Assessment (or Maple T.A.) improved students' results in different contexts. Pezzino [39] shows that the introduction of automatic formative assessment with Maple T.A. enabled interactive learning and visualization and improved students' results in an Advanced Economics course. Rønning [40] analyzed changes in the students' approach to problem solving after introducing Maple T.A. in Mathematics courses. At TU Delft, Möbius is successfully used to implement digital exams for Engineering courses [41].

Figure 1 shows an example of a question with automatic assessment created through Möbius Assessment. It asks students to write the equation of a line parallel to a given one and passing through a given point. The given line is displayed in a figure to help students visualize the problem. Students can visualize the graph of the answer they inserted, together with that of the given line, before submitting their answer, so they have the chance to reason more on their answer and modify it if they notice an incongruence. It is the first feedback given in a graphic register, while the answer is required in a symbolic register. The answer (an equation) is graded independently on the form thanks to the advanced computing environment running behind. At every student's attempt the coefficients of the given line and the y-coordinate of the given point change; the graph and the correct answer are computed through an algorithm and change accordingly.

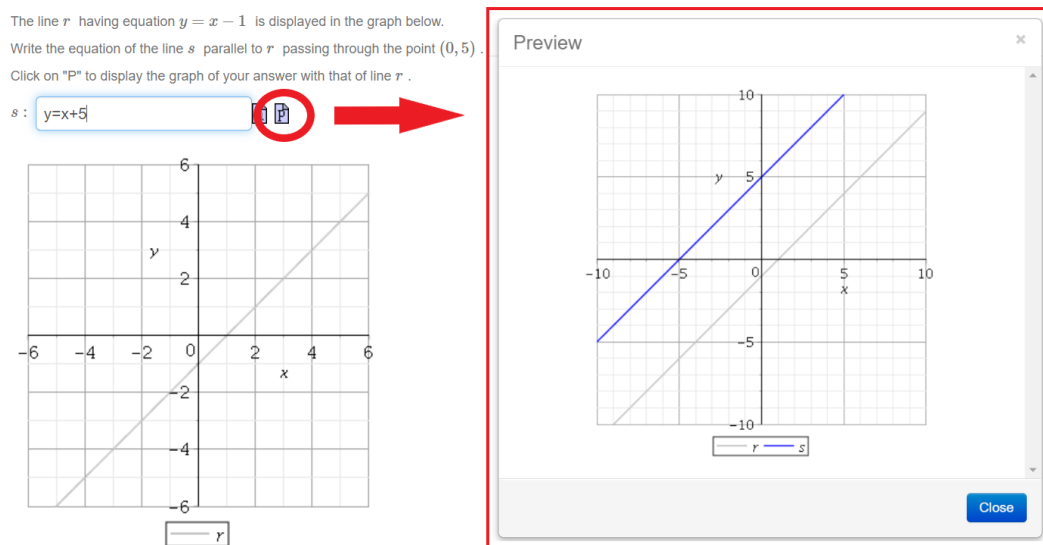


Figure 1. Example of a question created through Möbius Assessment.

2.4. Interactive Feedback and Mathematics

The results from the literature summarized above suggest that the feedback acquires effectiveness when it is student-centered, when it actively engages the student in an interactive dialogue, and when it elicits the students' actions to modify their learning trajectories [25–32]. Worked examples are commonly used as elaborated feedback and several studies show that they are more effective than corrective feedback [28,42]. However, they need to be carefully designed and properly used, otherwise students tend to skip them; moreover, they do not seem much effective to help students transfer the knowledge gained in the solving process to a new situation [43]. Digital technologies offer the possibility to create step-by-step interactive processes, which, in Mathematics, are commonly used to support students in problem-solving processes [44,45]. However, these step-by-step solving paths are usually not conceptualized as a feedback, but as a guidance to develop mathematical and problem-solving competences. In [45], Corbalan, Paas and Cuypers compare three different levels of formative feedback: feedback on the final solution step, feedback on all the solution steps at once and feedback on all the solution steps successively. Feedback on the final solution step is a sort of corrective feedback that does not give information about mistakes in the solving process but only on the final solution. Feedback on all the solution steps at once is a worked example used to compare one's solution with a correct one, while feedback on all the solution steps successively is given at each step of the solving process. The latter is implemented during a step-by-step guided problem-solving activity, but it leaves little freedom to the learners to invent a solving strategy, forcing them to follow the defined path. According to this study, the two kinds of feedback on all the solution steps are perceived as more useful than the feedback on the final solution only; the feedback on the solution steps successively seems the best one for more complex problems or novice learners. This study highlights the effectiveness of providing step-by-step worked examples; however, the interactive path as a whole is not conceptualized as a feedback. Rønning in [40] suggests, when using similar computerized step-by-step solutions, to be aware that they modify the problem itself and that students have to tackle a simpler problem than the original one.

Drawing on these results, we have developed the concept of “interactive feedback” in AFA activities as a step-by-step interactive solving process which shows a path to the solution after one or more autonomous attempt by the learner. It begins immediately after answering one question, when students are working on an online test. After showing the correctness of the answer, the system proposes a step-by-step resolution that interactively shows a possible process for solving the task. This interactive feedback can be displayed

only to the students who failed to answer autonomously to the main question, or even to those who made it correct. Its function can be guiding students in the argumentation of the solving process or comparing their solving strategy with a correct one. In the interactive feedback, sub-questions investigate prerequisites, simpler tasks, or other representations of the initial problem, to guide students to a possible way to tackle the task. At each step, if they give the wrong answer, the correct one is shown to be used in the following steps. Conceptualizing the step-by-step process as a feedback allows us to solve Rønning's concerns about giving hints on how to solve the problem: it happens in the feedback, not in the problem itself, and the sub-questions open to an interactive dialogue in which the learner is actively engaged. Our conceptualization of interactive feedback is a kind of elaborated feedback since it does not limit to correct the answer. It consists of a formative interaction between the students and the system, and actively engages students in resolving the task.

This schema is made possible by the “adaptive” modality available in Möbius Assessment; it can also be replicated with different automatic assessment systems having adaptive capabilities. Students can gradually acquire the background and the process that enables them to answer the initial problem. Moreover, they earn partial credits for the correctness of their answer in the step-by-step process. These points act as a motivational lever and, expressing intermediate levels between “incorrect” and “correct”, also offer teacher and students more precise information about the students' competence in a particular domain.

Interactive feedback can be particularly effective when it is inserted in the model for designing automatic formative assessment activities which we proposed in [46]. Besides interactive feedback, it includes:

- Availability of assignments, which can be attempted in a self-paced way, without limitations in data, duration, and number of attempts;
- Algorithm-based questions and answers, where random values, parameters, graphs, or formulas in the question text, answers, and feedback randomly change at every attempt and for each student;
- Open answers, especially mathematical ones, graded through the advanced computing capabilities of the system;
- Immediate feedback, shown to the students while they are still focused on the task; and
- Contextualization of the tasks in the real-world or interesting applications so that they can be relevant to students as well as for the discipline.

Figure 2 shows an example of a question with interactive feedback developed with Möbius Assessment. The question was conceived for students in grade 8 when working with linear models. It deals with a rocket that should dock at a spaceship, launched at a different time, and traveling at a different velocity. After the first section, where the final solution of the problem is asked, students are led through a step-by-step process showing different methods to tackle the problem: a numeric strategy (a table to be filled in) and a graphic strategy (four graphics among which to choose). In this way, students have the chance to explore the problem using two different representation registers. After this exploration, the two initial questions are posed again. At every attempt, an algorithm generates new data for the launch times and velocities, and updates graphics, formulas, and correct solutions accordingly.

Space Launch

An interplanetary rocket leaves the Earth atmosphere at 14 at the speed of 5000 km/h and it follows a linear trajectory with constant speed. The rocket must dock at spaceship which follows the same trajectory at the constant speed of 4000 km/h, which left the atmosphere 2 hours before, that is at 12.

At what time will the rocket dock at the spaceship?

At what distance from the atmosphere will they be?

Attempt 1 of 2

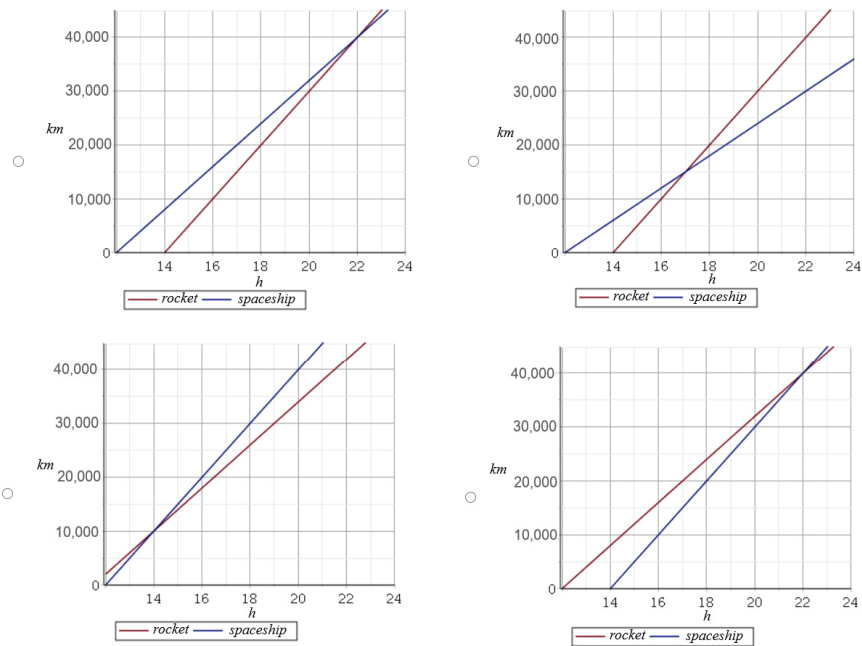


Complete the following table.

Time (h)	Distance from the atmosphere (km)	
	Spaceship	Rocket
12	0	0
13	<input type="text" value="4000"/> ✓	<input type="text" value="0"/> ✓
14	<input type="text" value="8000"/> ✓	<input type="text" value="0"/> ✓
15	<input type="text" value="12,000"/> ✓	<input type="text" value="5000"/> ✓
16	<input type="text" value="16,000"/> ✓	<input type="text" value="10,000"/> ✓
17	<input type="text" value="20,000"/> ✓	<input type="text" value="15,000"/> ✓
18	<input type="text" value="24,000"/> ✓	<input type="text" value="20,000"/> ✓



Choose the graphic representing the distances covered by rocket and spaceship.



At what time will the rocket dock at the spaceship? ✓



At what distance from the atmosphere will it happen? ✓

Figure 2. Example of a question with interactive feedback.

Many of this model’s features are typical of online assessment and used in many systems illustrated in the literature. The feedback’s immediacy is one of the main advantages of using automatic assessment acknowledged by many authors [47]. Regarding the timing of the feedback, there is disagreement in the definitions of immediate and delayed feedback

in the literature. While some authors define immediate feedback as the feedback provided after answering each item and delayed feedback like the one provided after a block of items, others include both cases in the “immediate feedback” and distinguish it from the “delayed feedback”, which is provided hours or even days after ending the test [28]. There is no doubt that immediate feedback is more useful than delayed feedback, intended as in the second definition; when considering the first definition’s distinction, there is disagreement about the best one. For instance, Fabienne van der Kleij and her colleagues, in a study with 153 university students, found that immediate feedback was more useful than delayed feedback for learning [27], while Gaona et al. [29] argue that delayed feedback is preferable to improve learning. In her review of the literature about feedback, Shute [28] argues that immediate feedback is more effective in the short term and for procedural knowledge, while delayed feedback may be superior for promoting the transfer of learning, especially concerning concept-formation tasks. Here we use feedback after each item when coupled with interactive feedback, which often has a procedural aim, and feedback at the end of the quiz when items do not include interactive feedback. Feedback can also be delayed by multiple attempts before showing the correct answer or the interactive feedback, to let students think more about the task before having the solution.

Availability is another relevant advantage of computer-based education and a key difference compared with paper-and-pen instruction. It refers to the possibility of finding learning materials independently of time and space and the chance to repeat the activities and receive timely feedback. This aspect, coupled with suitable activities, could help students study and learn at their own pace, supported continuously by the technologies [48,49]. Moreover, letting students repeat the assignments is an effective way to make them aware that information from the feedback was useful to improve and make teachers and researchers sure that feedback was well built [50].

In Mathematics Education, there are several examples of the use of algorithms to randomize parameters and of CAS-based engines to assess correct answers [29,36,51]. Random parameters open two scenarios: on one side, each student has a different version of the same item; on the other, each student can repeat the tests finding different questions, which maintain the same structure but have new numbers, or new formulas, graphs, or other elements each time. The algorithm can compute and generate even symbolic elements powered by a mathematical engine, thus allowing the creation of interesting items. Algorithmic questions have many advantages: the main ones are restrictions to cheating, saving time in the creation of large question banks, inserting every kind of mathematical object (e.g., graphs, formulas, matrixes, diagrams) in a question, and the possibility to write codes even to grade students’ answers [52,53].

The possibility to write, edit, and adapt grading-codes to one’s needs allows the assessment of open answers, through which open-ended tasks can be proposed, and question design becomes more flexible. Open-ended tasks are problems devoted to developing a mathematical idea that involves multiple methods, pathways, access points, and solutions [54]. They are powerful for creating opportunities for student exploration, collaboration, and sophisticated mathematical reasoning [55]; they permit students to access problems from different perspectives and evoke meaningful mathematical discourse [54]. Teaching through open-ended problems can foster students’ higher-order thinking, reflection, and enjoyment, besides offering students a greater opportunity to exercise metacognitive skills related to tool selection and decision making [56].

Contextualization of tasks in real-world settings contributes to creating meanings and a deeper understanding, as students can associate abstract concepts to real-life or concrete objects. It also works as a stimulus for motivation, since tasks can be closer to the students’ lives and interests [57].

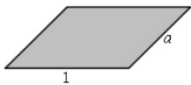
In a similar model for designing AFA activities, interactive feedback acquires even more importance. Students can try the initial problem on their own and, in case of a wrong answer, they have one or two more attempts available, so they are invited to focus more on the task and try again to reason on it. If they fail, they are shown the interactive feedback

which shows a possible approach to the solution with a step-by-step path. It can help them localize their mistakes or give a different idea to tackle the problem. Then, they can try the assignment again finding a similar problem, but with different data, so they have to repeat the whole process autonomously. Remembering the results by heart or developing a “trial and error” strategy will be useless. Thus, the interactive feedback is particularly relevant in making students process the feedback and use the information gained to improve their understanding. From a structural viewpoint, interactive feedback is part of the question itself, and the step-by-step process is shown immediately before moving on to the next question. Consequently, it should help ensure that students go through the feedback after coming to know if their answer was correct.

Interactive feedback can help low-level students because if they are discouraged by a complex task, they are engaged in simpler and more manageable sub-questions and avoid cognitive overload [43,45,56].

We would like to clarify here a point that will be later used in our analyses. Using an adaptive automatic assessment system such as Möbius Assessment, it is possible to create questions in a form of a step-by-step guided exercise or problem, which can be used to show a procedure or a paradigmatic example of problem-solving before asking them to autonomously solve tasks. We will call these kinds of questions “Guided Activities”. They are similar to the activities with “feedback on all the solution step successively” in Corbalan and colleagues’ study [45]. An example of a guided activity is shown in Figure 3. The activity is conceived for students in grade 8. It guides students in the exploration of a linear function generated by a geometry problem that is the perimeter of a parallelogram, where one side’s length is given through a variable. In the first step, the geometric situation is analyzed and students are asked to write a formula expressing the perimeter of the parallelogram. Notice that the formula is accepted independently of the form. In the second step, the formula is analyzed from a numeric point of view through a table to fill in. In the last step, students have to sketch the graph of the formula. Even if guided activities show interactive solving processes, they are not interactive feedback, for the simple reason that they are not feedback. In fact, the step-by-step process does not depend on an initial problem on which students have a chance to reason autonomously. They are very interesting for their design and they could be very effective to help students master processes or to have feedback about some results before using them for subsequent computations; however, we distinguish them from interactive feedback.

✓ Observe the parallelogram displayed in the following figure.



Write the formula expressing the perimeter of the parallelogram when a varies.

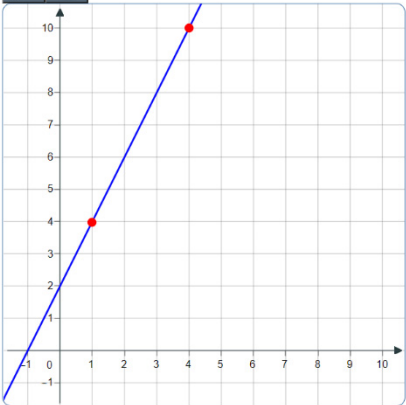
perimeter = ✓

Correct response: $2 + 2a$

✓ Fill in the table with the values of the perimeter corresponding to the given values of a .

a	Perimeter	
1	4	✓
2	6	✓
3	8	✓
4	10	✓

Now sketch the graph of the function perimeter = $2a + 2$



Attempt 1 of 1

Figure 3. Example of a guided activity.

3. Research Question and Method

In this paper, we consider Sadler's model of feedback as a reference for describing good feedback and formative assessment practices to which we aspire, and we investigate the following research questions:

(RQ1) Can the interactive feedback be effective for the development of Mathematical knowledge, according to Sadler's model [26]?

(RQ2) Is the improvement observed in activities with AFA influenced by the students' socio-economic background?

To investigate these research questions, we analyzed data of a bigger experimentation, named "Educating City", aimed at integrating innovative methodologies such as AFA and problem solving in a DLE and studying their impact to learn Mathematics at grade 8 [58]. The experimentation involved 399 students from six different lower secondary schools in the City of Turin (Italy) and their Mathematics teachers in the 2017/2018 school year.

Using data from national surveys conducted by INVALSI, the Italian institute in charge of evaluating the school system, and from the schools' self-evaluation reports, we grouped schools in two clusters based on the students' socio-economic status. The division broadly coincides with the schools' location: those located in the city center are mainly attended by medium-high socio-economic classes, while schools in the suburbs and mainly attended by students from lower socio-economic classes. Several activities were designed by Mathematics Education researchers according to the AFA model and a problem-solving approach and proposed to the classes. The activities could be used both in the classroom, working in groups with paper and pen, and the tasks displayed through the Interactive White Board, or at home, solving the tasks individually.

In this study, we consider the individual homework activities with AFA carried out by the students in the DLE. There were 68 assignments, all containing alternatively:

- One or two questions with interactive feedback, similar to that shown in Figure 2;
- One or two guided activities, similar to that shown in Figure 3;
- From one to five simple questions without adaptivity nor step-by-step solving processes, similar to that shown in Figure 1.

The class teachers gave students indications on what assignments to complete as homework, based on the topics covered during the lessons. They were free to repeat the assignments and they did not have any burden on what to complete.

To answer the research questions, we considered data about the platform usage, in particular, the number of attempts made by the students to each assignment and the grades obtained to the automatically assessed online tests during subsequent attempts, to understand if the interactive feedback helped students improve their performances. In detail, the total number of attempts to all the available tests—including repeated attempts to the same tests—was drawn from the platform for each student. We considered only the fully complete attempts, where the students requested and obtained feedback and grade. We excluded from the analysis the incomplete attempts, where the students just answered a few questions and left the system without grading the assignment nor viewing the feedback. We also excluded the cases where the students opened the assignments to look at the questions and did not answer them. This choice is in line with the models of formative assessment adopted in this paper. Without automatic grading and feedback, the activity would not have a formative value. Moreover, we considered the grades earned by the students in all the attempts, on a scale from 0 to 100.

We also classified the assignments into three groups:

- IF: assignments containing questions with interactive feedback;
- GA: assignments containing guided activities, exercises, or problems; and
- NA: non-adaptive assignments, without interactive feedback or guided exercises.

To evaluate whether the interactive feedback was effective according to Sadler's model, for each student and for each assignment group we computed:

- The average number of attempts per assignment;

- The average grade each student earned in their first assignment attempt;
- The average grade each student earned in their last assignment attempt; and
- The difference between the grade earned in the last attempt and the first attempt.

We also split the students' sample into two groups considering their socio-economic provenience.

Then we analyzed and cross-checked these data using SPSS 26.

4. Results

In the Educating City online course, there were 68 tests covering all the various topics usually studied in grade 8. They were classified based on the type of assignment, obtaining:

- 15 assignments contained questions with interactive feedback;
- 23 assignments contained guided exercises or problems; and
- 30 assignments included non-adaptive questions.

For each assignment, we computed the total number of students who attempted them at least once, the total number of attempts, and the ratio attempts/students. This ratio has 1 as a minimum value, since, for each assignment, we ignored students who did not make any attempt. We made this choice since teachers were free to use the assignments that they considered the most suitable for their lessons, so we do not expect that students have attempted all the assignments. Globally, each assignment was used by a minimum of 5 and a maximum of 167 students. The total number of attempts ranges from 7 to 497, and the ratio attempts/students ranges from 1 to 2.97. We added the type of assignment to this analysis: it emerged that the non-adaptive assignments registered the highest number of attempts, while the assignments with interactive feedback the lowest. We reported the results in Table 1. The ANOVA test on the three variables, considering the type of assignment as the independent variable, shows that the differences among the three categories are statistically significant (number of students: $F = 4.760$, critical value of $F: 3.138$, $p = 0.012$; number of attempts: $F = 4.778$, critical value of $F: 3.138$, $p = 0.012$; attempts/students: $F = 7.055$, critical value of $F: 3.138$, $p = 0.002$).

Table 1. Statistics on the attempts made for each assignment type.

Type		Number of Students	Number of Attempts	Attempts/Students
IF	Average	29.27	40.47	1.26
	SD	19.78	37.25	0.26
GA	Average	46.76	65.13	1.35
	SD	36.73	52.47	0.30
NA	Average	62.17	105.30	1.59
	SD	37.51	91.27	0.34
Total	Average	49.69	77.41	1.44
	SD	36.06	74.20	0.33
	<i>p</i> -value	0.012	0.012	0.002
	Sq. Eta	0.13	0.13	0.18

In addition, the Squared Eta test had significant results, showing that the assignment type explains the 13% of the variance of the number of students and assignments, and the 18% of the variance of the ratio attempts/students. We think that this distribution is due to an uneven distribution of the three types of assignments in the course sections. NA assignments were prevalent in some review sections, covering topics that students should have acquired in their background and that usually the teachers resume at the beginning of the school year. Many teachers told us to have assigned these topics for the Christmas holidays. Conversely, IF and GA assignments prevailed in the course sections mainly used towards the end of the school year and many teachers did not require those activities as mandatory.

As a second step, we considered the average grade each student earned in their first assignment attempt and the average grade each student earned in their last attempt. We grouped the assignments by type and computed, for each student, the average value per type of assignment (IF, GA, and NA). Then, we compared each couple of variables through pairwise Student's *t*-tests. We recall that, in questions with the interactive feedback, students earn the full grade if they answer correctly to the initial question. If they fail it, students are led to the interactive feedback, through which they can earn partial grades, usually up to a maximum of 80% of the question's full grade. Thus, by repeating the tasks, students have the chance to improve their scores.

We found that, for NA assignments, the average of the last grades remained similar to that of the first grades; for GA assignments, the difference is slightly higher but not outstanding, while for IF assignments, the improvement is considerable. All the increases, for the three couples of variables, are statistically significant to the Student's *t*-test (IF: $t = -4.153$, critical value of $t: 1.977$, $p < 0.001$; GA: $t = -3.136$, critical value of $t: 1.974$, $p = 0.002$; NA: $t = -2.001$, critical value of $t: 1.969$, $p = 0.047$), even if the most robust results are those concerning IF tests. In particular, the *p*-value of the test related to NA assignments is on the threshold of acceptability. Table 2 reports all the results. Here, N represents the total number of students who took at least one attempt at one assignment of that group.

Table 2. Analysis of the increases of the students' grades through their attempts at the assignments, grouped by type.

Type	N		First Attempt	Last Attempt	Increase
IF	138	Average	43.40	52.51	8.86
		SD	26.49	29.85	25.46
GA	170	Average	38.55	41.02	2.45
		SD	23.31	23.98	10.21
NA	251	Average	43.77	45.58	1.81
		SD	21.71	23.68	14.33

For each student, we also computed the difference between the score obtained during the last attempt and the first attempt to each assignment type, measuring the increase in their competence level during subsequent attempts. Results are listed in Table 2. We ran the ANOVA test to compare the average of the increases in the three groups of assignments. We found out that the means of the three groups are significantly different ($F = 8.64$, critical value of $F: 3.01$, $p < 0.001$), in particular, the increase for IF assignments is particularly higher than the other two). Thus, we can affirm that interactive feedback is more effective than other kinds of feedback in improving students' performance. It is interesting to notice that we obtained these results even if students tended to repeat more non-adaptive assignments or guided activities than those with IF. However, repeating assignments without IF, their results registered lower increases. The average score growth achieved through IF is about one point out of ten: it is a respectable value that proves an improvement in the students' competences.

We would like to point out that, in the analyses presented above, we did not consider if students actually made more attempts to the same assignments, but we included in the analyses also those students who made one only attempt at an assignment, so their grade in the last attempt is the same than the first attempt. The decision to include these data is in line with the conception of formative assessment as a practice that should generate some action aimed at covering the gap between current and reference performance. If students did not repeat the assignment, they missed a chance to improve, so their increase remained zero.

To measure the net effect of repeating the assignments of the three types on learning, we selected, for each group of assignments, only the students who made at least two attempts to at least one assignment of the group and repeated the analyses on the average grades of first and last attempts. In this case, the sample was reduced considerably; however, numbers are sufficiently high to perform statistical analyses. We found that,

for non-adaptive assignments, the average grade increased by 2.70 points out of 100; for guided activities, it increased by 4.36 points out of 10; for assignments with interactive feedback, the registered growth consisted in 22.07 points out of 100. While for NA and GA assignments the increase remained under 5 points out of 100, the increase for IF assignments was remarkable and exceeded 20% of the total points. Results are listed in Table 3. As done before, we ran the ANOVA test to compare the average of the increases in the three groups of assignments. We found out that the means of the three groups are still more significantly different ($F = 18.58$, critical value of $F: 3.02$, $p < 0.001$); the increase for IF assignments clearly stands out among the others.

Table 3. Analysis of the increases of the students' grades after multiple attempts at the assignments, grouped by type.

Type	N		First Attempt	Last Attempt	Increase
IF	57	Average	35.88	57.95	22.07
		SD	25.99	34.30	36.23
GA	96	Average	39.65	44.02	4.36
		SD	21.05	21.95	13.28
NA	168	Average	47.01	49.72	2.70
		SD	18.79	21.59	17.41

Comparing the results in Table 3 with those in Table 2, we can notice that the first attempt in IF assignments in Table 3 is remarkably lower than the same value in Table 2, while the other means are rather comparable (first attempts in GA and NA assignments are slightly higher in Table 3 than Table 2). This means that students who made multiple attempts in IF assignments earned lower marks in their first attempt than those who made one only attempt. In other words, students who earned low marks in IF assignments (and went through the interactive feedback) tended to repeat the assignment, while this tendency is not registered in the other assignment types. This result is relevant for RQ1, since it shows that the IF engages students who most need to close a gap between their actual level and the desired one in actions to improve their results, namely, repeating the assignments. Assignments without interactive feedback are repeated similarly by students of lower and higher levels.

Lastly, we focused on (RQ2) and investigated if there was a relation between the increase in the three kinds of assignments' scores after multiple attempts and the students' socio-economic class. Thus, we ran three ANOVA tests on the three variables registering the grade increase in IF, GA, and NA assignments, using the socio-economic level as the independent variable. Results are shown in Table 4. For all the three assignment types, the increase registered in lower classes are higher than medium-high classes. For NA assignments the difference is little and not significant ($F = 2.218$, critical value of $F: 3.898$, $p = 0.138$), and both the increases are quite little. Conversely, for GA and IF assignments, the difference is higher and significant (GA: $F = 18.109$, critical value of $F: 3.942$, $p < 0.001$; IF: $F = 9.104$, critical value of $F: 4.016$, $p = 0.004$). In particular, the increase in IF activities for students from disadvantaged backgrounds is a very high value, namely, 36.46 out of 100. It means that the students from lower social classes participating in this project who repeated an IF assignment at least twice improved their grades on average of 36%. Moreover, we can notice that the increase in GA for students from medium-high socio-economic contexts is about zero: it seems that GA did not help students from wealthier families improve their results, while they had a medium effect on students from disadvantaged backgrounds.

Table 4. Analysis of the increases of the students' grades after multiple attempts to the assignments, by socio-economic level.

Socio-Economic Status		Increase (IF)	Increase (GA)	Increase (NA)
Low	Average	36.46	11.25	4.66
	SD	40.44	15.72	12.51
	N	27	36	86
Medium-high	Average	9.12	0.23	0.66
	SD	27.32	9.70	21.36
	N	30	60	83

To conclude our analysis, we examined the answers registered in the DLE to IF assignments in multiple attempts and we chose a significant example. Emily (the name is imaginary), a student from a challenging background and second-generation immigrant, was working on the "Space Launch" task shown in Figure 2. She made three attempts and each time she found different values in the problem's text. Her first attempt is shown in Figure 4. She failed the answer (it seems she tried to guess) so she went through the interactive feedback. She filled in the table understanding how it should be completed but making a mistake in the rocket's position at its launch time. She received as feedback the table's correct values and she used them to choose the graph. Here she made another mistake, inverting the two functions. As feedback, she got the correct graph and she correctly used this information in the last two parts, where she identified the correct intersection point. Her score was 30/100. She immediately ran a second attempt, finding different values. Now she correctly found the docking time but not the distance, so she went through the interactive feedback. Now she correctly completed the table (demonstrating having understood her previous mistake) but still repeated the mistake in the graph's choice. She correctly completed the last two questions. Her score was 55/100. She immediately ran the third attempt and this time she correctly answered the initial problem, earning the maximum score without visualizing the interactive feedback.

Through this episode, we can see how the interactive feedback helped Emily to understand a solving strategy and apply it to solve the problem. The IF engaged the student in attempting the task again until she mastered the solving process. It also helped her identify her mistakes, which she self-corrected in subsequent attempts. Her grade increased to 70 points out of 100.

Emily's response	Correct response
Space Launch	Space Launch

An interplanetary rocket leaves the Earth atmosphere at 17 at the speed of 8000 km/h and it follows a linear trajectory with constant speed. The rocket must dock at spaceship which follows the same trajectory at the constant speed of 4000 km/h, which left the atmosphere 2 hours before, that is at 15.

At what time will the rocket dock at the spaceship?

✗ 18

✗ 18

At what distance from the atmosphere will they be?

✗ 21

✗ 21

An interplanetary rocket leaves the Earth atmosphere at 17 at the speed of 8000 km/h and it follows a linear trajectory with constant speed. The rocket must dock at spaceship which follows the same trajectory at the constant speed of 4000 km/h, which left the atmosphere 2 hours before, that is at 15.

At what time will the rocket dock at the spaceship? 19

At what distance from the atmosphere will they be? 16,000 km



Complete the following table.

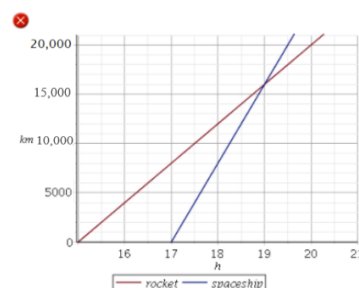
Time (h)	Distance from the atmosphere (km)	
	Spaceship	Rocket
15	0	0
16	✓ 4000	✓ 0
	✓ 4000	✓ 0
17	✓ 8000	✗ 8000
	✓ 8000	✗ 8000
18	✓ 12,000	✗ 16,000
	✓ 12,000	✗ 16,000
19	✓ 16,000	✗ 24,000
	✓ 16,000	✗ 24,000
20	✓ 20,000	✗ 32,000
	✓ 20,000	✗ 32,000
21	✓ 24,000	✗ 40,000
	✓ 24,000	✗ 40,000

Complete the following table.

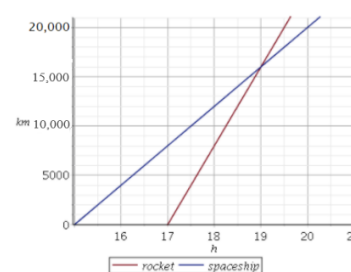
Time (h)	Distance from the atmosphere (km)	
	Spaceship	Rocket
15	0	0
16	4000	0
17	8000	0
18	12,000	8000
19	16,000	16,000
20	20,000	24,000
21	24,000	32,000



Choose the graphic representing the distances covered by rocket and spaceship.



Choose the graphic representing the distances covered by rocket and spaceship.



At what time will the rocket dock at the spaceship? 19

✓

At what time will the rocket dock at the spaceship? 19



At what distance from the atmosphere will it happen? 16,000 km

✓

At what distance from the atmosphere will it happen? 16,000 km



Figure 4. Emily's answers to the "Space Launch" activity at her first attempt.

5. Discussions and Conclusions

In this paper, we have presented and discussed our conceptualization of interactive feedback for AFA in Mathematics, as a step-by-step interactive solving process which guides the learner in solving a task after one or more autonomous attempts. We grounded our research on studies on effective feedback for closing the gap between actual and desired performance. The IF can be proposed as a kind of elaborated feedback [25]; it differs from the most typical kind of elaborated feedback, namely, worked example, since it is interactive; it is also different from giving hints or checking all the solving steps [42,43] during the students' work, since the IF comes after the autonomous attempt, it is a feedback which encourage learners to compare their solving strategy with a given one in an interactive way. We integrated the IF in a wider model for the design of AFA activities, mainly based on the following other features: availability, algorithm-based questions and answers, open-ended answers, immediate feedback, and real-world contextualization.

Through an experimentation involving a total number of 299 Italian students in grade 8, we showed that the IF can be useful for ensuring that students use the feedback information to improve their performance, reducing the gap between current and reference performance.

In particular, the results helped us answer (RQ1), "Can the interactive feedback be effective for the development of Mathematical knowledge, according to Sadler's model [26]?" In fact:

- a. It offers students a concept of the standard that they can actively possess, through an example of the correct resolution of the task;
- b. It helps students compare the actual (or current) level of performance with the standard through the interactive process and the immediate feedback of the sub-tasks; and
- c. It engages students in an appropriate action, which leads to some closure of the gap: in fact, students who achieved low grades tended to repeat the task, finding a similar task with different values, so they had to repeat the process. The remarkable improvement in their results shows that the interactive feedback helped them master the solving process.

Thus, we have positively answered (RQ1): the interactive feedback can be effective for the development of Mathematical knowledge, according to Sadler's model.

We could also answer (RQ2), "Is the improvement observed in activities with AFA influenced by the students' socio-economic background?" The results of the analyses show that IF activities noticeably helped students from lower social classes improve their results when reattempting the assignments, and significantly more than those from wealthier backgrounds. Additionally, guided activities were more effective for students from lower socio-economic classes than the others, while the socio-economic provenience seems not to influence the improvements in activities with only corrective feedback and without adaptive features. Therefore, we have an answer for (RQ2): the socio-economic background influenced the students' improvements in adaptive activities (IF and GA assignments); in particular, students from lower socio-economic classes improved significantly more. Conversely, there was no influence of the socio-economic classes on the improvement in non-adaptive activities.

We are aware that the numbers involved in this experimentation are limited, and the number of students who made repeated attempts at the assignments was quite low, so we cannot generalize these results. However, these results can encourage us to develop further the research about interactive feedback and adaptive activities and inspire new research paths.

Our model has been designed and experimented before the COVID-19 pandemic, so it has been conceived to be used to support face-to-face, blended, or online education. In particular, the experimentation object of this study was held in a blended modality, and IF activities were mainly proposed as homework, with the idea of letting students be guided by the IF in case of difficulties. During focus groups with teachers, it emerged

that the IF resolved a great part of the difficulties, so the time usually spent discussing and correct the homework could be spent in other learning activities. We believe that this kind of feedback would be quite valuable in a distance learning context, such as during the current health emergency, since it helps students self-assess, autonomously individuate and correct their mistakes, and master solving strategies. In a word, it can help students develop self-regulation competence, a precious competence for these times and their future. We are working on other research projects aimed to study the effectiveness of the IF in the pandemic context.

IF activities can be developed and adapted for all grades, from primary school to university. Currently, we are studying the effect of IF at different stages of education, to compare its impact and effectiveness on learning.

This study can inspire the research about task design for formative assessment, as well as teachers, educators, and content authors. The results shown above suggest that adaptive activities could be more effective than more traditional ones to help students covering the gap between current and desired knowledge and competence level. Among the adaptive activities, those structured as interactive feedback resulted in the most effective for this purpose with students in grade 8, while guided activities had a significant but minor effect. It would be interesting to understand if these results would depend on the students' age and expand the research in this direction. Moreover, guided activities could be effective for different purposes, such as supporting students in understanding new processes, simple demonstrations, or deductive reasoning.

A very interesting result obtained with this study is the particular effectiveness of IF for students from disadvantaged backgrounds. Thus, interactive feedback can be extremely useful during (and after) the health emergency, since it has been observed that students from lower social classes have the highest risk of registering important learning losses in Mathematics [8]. IF could provide them with a guide and motivation to tackle the difficulties. These results are in line with other results obtained by our research group in this project: we had previously shown that interactive activities with AFA in a DLE had a positive impact on the improvement of the levels of school engagement on students from lower socio-economic classes [58,59]. We believe that the reason can be situated exactly in the interactive feedback, which powerfully engages the learners in an interactive dialogue with the technology which activates cognitive processes. It manages to generate internal feedback, which helps students understand and monitor their learning, so it is an "interactive feedback", as Nicol also ascertains [32]. From these results, we can infer some recommendations for policymakers interested in improving the quality of learning for students from lower socio-economic classes by acting on engagement, or, more generally, for students with difficulties in Mathematics. Supporting projects and actions based on Automatic Formative Assessment activities with interactive feedback could help young students engage with Mathematics and develop important competences for their future [10].

Interactive feedback is a relevant formative assessment practice, which enables the collection of precise data on the students' level of competence and their progress. When students solve a task with IF, a moment of contingency, as theorized by Wiliam [19], is generated: the structure of the activity changes based on the answer, the need of examining in-depth the problem is evidenced and the guided path starts. Teachers receive clear information about the students' difficulties and their need to revise a specific topic or process and they can integrate the IF with other activities. This consideration gives prompts to the research in teacher training: for efficient use of AFA activities, teachers should be adequately prepared to read and understand the information gained from automatic assessment and to use it to make decisions. Moreover, they should also be trained to use an automatic assessment system autonomously, and even to create their own materials, so that experiences such as the one described in this paper could spread and become daily practices. All the materials used in the experimentation presented here are freely available to all the Italian teachers through the Problem Posing and Solving Project of the Ministry of

Education [60]. It is a project aimed at innovating the teaching and learning of Mathematics and other scientific disciplines in secondary schools through teacher training and sharing good practices in a DLE. We have also been involved in several training courses on AFA.

IF activities are a starting point to develop adaptive activities, which ask students to solve different tasks based on their proficiency level [61]. For the development of this theory, it would be interesting to use data from IF to identify difficulties and misunderstandings in Mathematics to drive students towards focused and appropriate activities, also based on the students' attitudes and learning styles. A similar study would move in the Learning Analytics direction and would aim at developing solutions able to automatically detect difficulties and help teachers make decisions to dynamically shape the learning path [4,62,63].

Author Contributions: Conceptualization, A.B., M.M. and M.S.; methodology, A.B., M.M. and M.S.; formal analysis, A.B., M.M. and M.S.; investigation, A.B., M.M. and M.S.; data curation, A.B., M.M. and M.S.; writing—original draft preparation, A.B., M.M. and M.S.; writing—review and editing, A.B., M.M. and M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received funds from the Compagnia di San Paolo Banking Foundation through the “OPERA—Open Program for Educational Resources and Activities” and “ExPost” Projects.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy reasons.

Acknowledgments: The authors would like to thank the Compagnia di San Paolo for the financial support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yerushalmy, M.; Nagari-Haddif, G.; Olsher, S. Design of Tasks for Online Assessment that Supports Understanding of Students' Conceptions. *ZDM* **2017**, *49*, 701–716. [\[CrossRef\]](#)
2. Stacey, K.; Wiliam, D. Technology and Assessment in Mathematics. In *Third International Handbook of Mathematics Education*; Clements, M.A., Ed.; Springer International Handbooks of Education; Springer: New York, NY, USA, 2013; Volume 27, pp. 721–751.
3. Moreno, J.; Pineda, A.F. A Framework for Automated Formative Assessment in Mathematics Courses. *IEEE Access* **2020**, *8*, 30152–30159. [\[CrossRef\]](#)
4. Barana, A.; Conte, A.; Fissore, C.; Marchisio, M.; Rabellino, S. Learning Analytics to Improve Formative Assessment Strategies. *J. e-Learn. Knowl. Soc.* **2019**, *15*, 75–88. [\[CrossRef\]](#)
5. Bolondi, G.; Del Zozzo, A.; Ferretti, F.; Garzetti, M.; Santi, G. Can Formative Assessment Practices Appear Spontaneously during Long Distance Learning? In *CSEU 2021, Proceedings of the 13th International Conference on Computer Supported Education, Online, 23–25 April 2021*; Csapó, B., Uhomoihi, J., Eds.; SCITEPRESS: Setúbal, Portugal, 2021; Volume 1, pp. 625–632.
6. Darling-Hammond, L.; Schachner, A.; Edgerton, A.K. *Restarting and Reinventing School: Learning in the Time of COVID and Beyond*; Learning Policy Institute: Palo Alto, CA, USA, 2020.
7. Meeter, M. Primary School Mathematics during Covid-19: No Evidence of Learning Gaps in Adaptive Practicing Results. *PsyArXiv* **2021**. [\[CrossRef\]](#)
8. Kuhfeld, M.; Soland, J.; Tarasawa, B.; Johnson, A.; Ruzek, E.; Liu, J. Projecting the Potential Impact of COVID-19 School Closures on Academic Achievement. *Educ. Res.* **2020**, *49*, 549–565. [\[CrossRef\]](#)
9. Dorn, E.; Hancock, B.; Sarakatsannis, J.; Viruleg, E. *COVID-19 and Learning Loss—Disparities Grow and Students Need Help*; McKinsey & Company: New York, NY, USA, 2020.
10. European Union. *European Parliament and Council Recommendation of 22 May 2018 on Key Competences for Lifelong Learning*; Official Journal of the European Union; European Union: Brussels, Belgium, 2018; pp. 1–13.
11. 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*; Taylor and Francis Group Ltd.: London, UK, 2006; pp. 1–14.
12. Nicol, D.J.; Macfarlane-Dick, D. Formative Assessment and Self-regulated Learning: A Model and Seven Principles of Good Feedback Practice. *Stud. High. Educ.* **2006**, *31*, 199–218. [\[CrossRef\]](#)
13. Moreno, R.; Mayer, R. Interactive Multimodal Learning Environments. *Educ. Psychol. Rev.* **2007**, *19*, 309–326. [\[CrossRef\]](#)

14. Fissore, C.; Marchisio, M.; Rabellino, S. Secondary School Teacher Support and Training for Online Teaching during the Covid-19 Pandemic. In *Proceedings of the European Distance and E-Learning Network (EDEN) Proceedings, Online, 22–24 June 2020*; European Distance and E-Learning Network: Timisoara, Romania, 2020; pp. 311–320.
15. Barana, A.; Marchisio, M. Analyzing Interactions in Automatic Formative Assessment Activities for Mathematics in Digital Learning Environments. In *CSEDU 2021, Proceedings of the 13th International Conference on Computer Supported Education, Online, 23–25 April 2021*; Csapó, B., Uhomobhi, J., Eds.; SCITEPRESS: Setúbal, Portugal, 2021; Volume 1, pp. 497–504. [\[CrossRef\]](#)
16. Jaques, P.A.; Seffrin, H.; Rubi, G.; de Morais, F.; Ghilardi, C.; Bittencourt, I.I.; Isotani, S. Rule-Based Expert Systems to Support Step-by-Step Guidance in Algebraic Problem Solving: The Case of the Tutor PAT2Math. *Expert Syst. Appl.* **2013**, *40*, 5456–5465. [\[CrossRef\]](#)
17. Kilicman, A.; Hassan, M.A.; Husain, S.K.S. Teaching and Learning Using Mathematics Software “The New Challenge”. *Procedia Soc. Behav. Sci.* **2010**, *8*, 613–619. [\[CrossRef\]](#)
18. Marchisio, M.; Remogna, S.; Roman, F.; Sacchet, M. Teaching Mathematics in Scientific Bachelor Degrees Using a Blended Approach. In *Proceedings of the 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC 2020)*, Madrid, Spain, 13–17 July 2020; pp. 190–195. [\[CrossRef\]](#)
19. Black, P.; Wiliam, D. Developing the Theory of Formative Assessment. *Educ. Assess. Eval. Account.* **2009**, *21*, 5–31. [\[CrossRef\]](#)
20. Wiliam, D. Formative Assessment: Getting the Focus Right. *Educ. Assess.* **2006**, *11*, 283–289. [\[CrossRef\]](#)
21. Hattie, J.A.C. *Visible Learning: A Synthesis of over 800 Meta-Analyses Relating to Achievement*; Routledge: London, UK, 2009; ISBN 0-203-88733-6.
22. Azevedo, R.; Bernard, R. A Meta-Analysis of the Effects of Feedback in Computer-Based Instruction. *J. Educ. Comput. Res.* **1995**, *13*, 111–127. [\[CrossRef\]](#)
23. Kluger, A.N.; DeNisi, A. The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory. *Psychol. Bull.* **1996**, *119*, 254–284. [\[CrossRef\]](#)
24. Hattie, J.; Timperley, H. The Power of Feedback. *Rev. Educ. Res.* **2007**, *77*, 81–112. [\[CrossRef\]](#)
25. Timmers, C.; Veldkamp, B. Attention Paid to Feedback Provided by a Computer-Based Assessment for Learning on Information Literacy. *Comput. Educ.* **2011**, *56*, 923–930. [\[CrossRef\]](#)
26. Sadler, D.R. Formative Assessment and the Design of Instructional Systems. *Instr. Sci.* **1989**, *18*, 119–144. [\[CrossRef\]](#)
27. van der Kleij, F.M.; Eggen, T.J.H.M.; Timmers, C.F.; Veldkamp, B.P. Effects of Feedback in a Computer-Based Assessment for Learning. *Comput. Educ.* **2012**, *58*, 263–272. [\[CrossRef\]](#)
28. Shute, V.J. Focus on Formative Feedback. *Rev. Educ. Res.* **2008**, *78*, 153–189. [\[CrossRef\]](#)
29. Gaona, J.; Reguant, M.; Valdivia, I.; Vásquez, M.; Sancho-Vinuesa, T. Feedback by Automatic Assessment Systems Used in Mathematics Homework in the Engineering Field. *Comput. Appl. Eng. Educ.* **2018**, *26*, 994–1007. [\[CrossRef\]](#)
30. Chung, G.K.W.K.; Shel, T.; Kaiser, W.J. An Exploratory Study of a Novel Online Formative Assessment and Instructional Tool to Promote Students’ Circuit Problem Solving. *J. Technol. Learn. Assess.* **2006**, *5*, 4–26.
31. Black, P.; Wiliam, D. ‘In Praise of Educational Research’: Formative Assessment. *Br. Educ. Res. J.* **2003**, *29*, 623–637. [\[CrossRef\]](#)
32. Nicol, D. Reconceptualising Feedback as an Internal Not an External Process. *Ital. J. Educ. Res.* **2019**, *12*, 71–83.
33. van den Bogaard, M.E.D.; Saunders-Smits, G.N. Peer & Self Evaluations as Means to Improve the Assessment of Project Based Learning. In *Proceedings of the 2007 37th Annual Frontiers in Education Conference-Global Engineering: Knowledge without Borders, Opportunities without Passports*, Milwaukee, WI, USA, 10–13 October 2007; Volume Session S1G, pp. 12–18.
34. Gikandi, J.W.; Morrow, D.; Davis, N.E. Online Formative Assessment in Higher Education: A Review of the Literature. *Comput. Educ.* **2011**, *57*, 2333–2351. [\[CrossRef\]](#)
35. Hallam, G.; Glanville, C. Blending Process with Product: Using Assessment to Drive Learning Through the Creation of an Online Journal. In *Enhancing Teaching and Learning through Assessment: Deriving an Appropriate Model*; Frankland, S., Ed.; Springer: Dordrecht, The Netherlands, 2007; pp. 242–255. ISBN 978-1-4020-6225-4.
36. Sangwin, C. Computer Aided Assessment of Mathematics Using STACK. In *Selected Regular Lectures from the 12th International Congress on Mathematical Education*; Cho, S.J., Ed.; Springer: Cham, Switzerland, 2015; pp. 695–713. ISBN 978-3-319-17186-9.
37. Barana, A.; Marchisio, M.; Sacchet, M. Advantages of Using Automatic Formative Assessment for Learning Mathematics. In *Technology Enhanced Assessment*; Draaijer, S., Joosten-ten Brinke, D., Ras, E., Eds.; Springer: Cham, Switzerland, 2019; Volume 1014, pp. 180–198. ISBN 978-3-030-25263-2. [\[CrossRef\]](#)
38. 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.* **1989**, *20*, 561–569. [\[CrossRef\]](#)
39. Pezzino, M. Online Assessment, Adaptive Feedback and the Importance of Visual Learning for Students. The Advantages, with a Few Caveats, of Using MapleTA. *Int. Rev. Econ. Educ.* **2018**, *28*, 11–28. [\[CrossRef\]](#)
40. Rønning, F. Influence of Computer-Aided Assessment on Ways of Working with Mathematics. *Teach. Math. Its Appl. Int. J. IMA* **2017**, *36*, 94–107. [\[CrossRef\]](#)
41. Keijzer-de Ruijter, M. Digital Exams in Engineering Education. In *Technology Enhanced Assessment. TEA 2018*; Draaijer, S., Joosten-ten Brinke, D., Eds.; Communications in Computer and Information Science; Springer: Cham, Switzerland, 2019; Volume 1014.
42. Van der Kleij, F.M.; Lipnevich, A.A. Student Perceptions of Assessment Feedback: A Critical Scoping Review and Call for Research. *Educ. Assess. Eval. Acc.* **2021**, *33*, 345–373. [\[CrossRef\]](#)

43. Manson, E.; Ayres, P. Investigating How Errors Should Be Flagged and Worked Examples Structured When Providing Feedback to Novice Learners of Mathematics. *Educ. Psychol.* **2021**, *41*, 153–171. [[CrossRef](#)]
44. Beevers, C.E.; Paterson, J.S. Automatic Assessment of Problem Solving Skills in Mathematics. *Act. Learn. High. Educ.* **2003**, *4*, 127–144. [[CrossRef](#)]
45. Corbalan, G.; Paas, F.; Cuypers, H. Computer-Based Feedback in Linear Algebra: Effects on Transfer Performance and Motivation. *Comput. Educ.* **2010**, *55*, 692–703. [[CrossRef](#)]
46. Barana, A.; Conte, A.; Fioravera, M.; Marchisio, M.; Rabellino, S. A Model of Formative Automatic Assessment and Interactive Feedback for STEM. In Proceedings of the 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), Tokyo, Japan, 23–27 July 2018; pp. 1016–1025. [[CrossRef](#)]
47. Kurvinen, E.; Lindén, R.; Rajala, T.; Kaila, E.; Laakso, M.-J.; Salakoski, T. Automatic Assessment and Immediate Feedback in First Grade Mathematics. In Proceedings of the 14th Koli Calling International Conference on Computing Education Research—Koli Calling'14, Koli, Finland, 20–23 November 2014; ACM Press: Koli, Finland, 2014; pp. 15–23.
48. Barana, A.; Marchisio, M.; Miori, R. MATE-BOOSTER: Design of Tasks for Automatic Formative Assessment to Boost Mathematical Competence. In *Computer Supported Education*; Lane, H.C., Zvacek, S., Uhomobhi, J., Eds.; Communications in Computer and Information Science; Springer International Publishing: Cham, Switzerland, 2020; Volume 1220, pp. 418–441. ISBN 978-3-030-58458-0. [[CrossRef](#)]
49. Bruschi, B.; Cantino, V.; Cavallo Perin, R.; Culasso, F.; Giors, B.; Marchisio, M.; Marelli, C.; Milani, M.; Operti, L.; Parola, A.; et al. Start@unito: A Supporting Model for High School Students Enrolling to University. In Proceedings of the 15th International Conference CELDA 2018: Cognition and Exploratory Learning in Digital Age, Budapest, Hungary, 21–23 October 2018; pp. 307–312.
50. Boud, D. Sustainable Assessment: Rethinking Assessment for the Learning Society. *Stud. Contin. Educ.* **2000**, *22*, 151–167. [[CrossRef](#)]
51. Cazes, C.; Gueudet, G.; Hersant, M.; Vandebrouck, F. Using E-Exercise Bases in Mathematics: Case Studies at University. *Int. J. Comput. Math. Learn.* **2007**, *11*, 327–350. [[CrossRef](#)]
52. Barana, A.; Fissore, C.; Marchisio, M. From Standardized Assessment to Automatic Formative Assessment for Adaptive Teaching. In Proceedings of the 12th International Conference on Computer Supported Education, Online, 2–4 May 2020; SCITEPRESS-Science and Technology Publications: Prague, Czech Republic, 2020; pp. 285–296. [[CrossRef](#)]
53. Marchisio, M.; Margaria, T.; Sacchet, M. Automatic Formative Assessment in Computer Science: Guidance to Model-Driven Design. In Proceedings of the 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC), Madrid, Spain, 13–17 July 2020; pp. 201–206. [[CrossRef](#)]
54. Edson, A.J. Learner-Controlled Scaffolding Linked to Open-Ended Problems in a Digital Learning Environment. *ZDM* **2017**, *49*, 735–753. [[CrossRef](#)]
55. Chan, M.C.E.; Clarke, D. Structured Affordances in the Use of Open-Ended Tasks to Facilitate Collaborative Problem Solving. *ZDM* **2017**, *49*, 951–963. [[CrossRef](#)]
56. Boaler, J. Open and Closed Mathematics: Student Experiences and Understandings. *J. Res. Math. Educ.* **1998**, *29*, 41–62. [[CrossRef](#)]
57. Samo, D.D.; Darhim, D.; Kartasasmita, B. Culture-Based Contextual Learning to Increase Problem-Solving Ability of First Year University Student. *J. Math. Educ.* **2017**, *9*. [[CrossRef](#)]
58. Barana, A.; Boffo, S.; Gagliardi, F.; Garuti, R.; Marchisio, M. Empowering Engagement in a Technology-Enhanced Learning Environment. In *Project and Design Literacy as Cornerstones of Smart Education*; Rehm, M., Saldien, J., Manca, S., Eds.; Springer: Singapore, 2020; Volume 158, pp. 75–77. ISBN 9789811396519.
59. Barana, A.; Marchisio, M. An Interactive Learning Environment to Empower Engagement in Mathematics. *IxD&A* **2020**, *45*, 302–321.
60. Fissore, C.; Floris, F.; Marchisio, M.; Rabellino, S.; Sacchet, M. Digital Competences for Educators in the Italian Secondary School: A Comparison between DigCompEdu Reference Framework and the PP&S Project Experience. In Proceedings of the International Conference E-Learning 2020, Online, 21–23 July 2020; IADIS Press: Rome, Italy, 2020; pp. 47–54. [[CrossRef](#)]
61. Di Caro, L.; Fioravera, M.; Marchisio, M.; Rabellino, S. Towards Adaptive Systems for Automatic Formative Assessment in Virtual Learning Communities. In Proceedings of the 2018 IEEE 42nd Annual Computer Software and Applications Conference COMPSAC, Tokyo, Japan, 23–27 July 2018; pp. 1000–1005. [[CrossRef](#)]
62. Tempelaar, D.T.; Heck, A.; Cuypers, H.; van der Kooij, H.; van de Vrie, E. Formative Assessment and Learning Analytics. In Proceedings of the Third International Conference on Learning Analytics and Knowledge—LAK'13, Leuven, Belgium, 8–12 April 2013; ACM Press: Leuven, Belgium, 2013; p. 205.
63. Marchisio, M.; Rabellino, S.; Roman, F.; Sacchet, M.; Salusso, D. Boosting Up Data Collection and Analysis to Learning Analytics in Open Online Contexts: An Assessment Methodology. *J. e-Learn. Knowl. Soc.* **2019**, *15*, 11. [[CrossRef](#)]