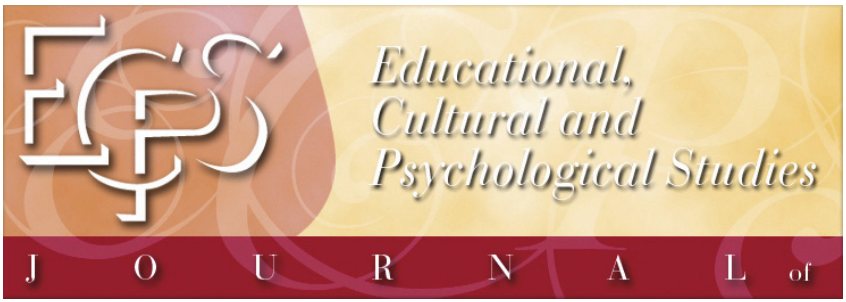


31
June 2025

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Note di Ricerca
Research Notes

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Formative Assessment Practices for Improving Students' Text Comprehension Abilities: An Experiment in a Lower Secondary School in Italy

Elisa Guasconi - Ira Vannini *

Alma Mater Studiorum Università di Bologna - Department of Education Studies «Giovanni Maria Bertin» (Italy)

DOI: https://doi.org/10.7358/ecps-2025-031-guva

elisa.guasconi2@unibo.it
ira.vannini@unibo.it

PRASSI DI «FORMATIVE ASSESSMENT» PER PROMUOVERE LE ABILITÀ DI COMPRENSIONE DEL TESTO: UNA SPERIMENTAZIONE NELLA SCUOLA SECONDARIA DI PRIMO GRADO IN ITALIA

Abstract

Formative assessment (FA) is an expression that covers several practices. While this variety has led to the development of different studies in the international context, there is a lack of experimental research in Italy on the topic. Therefore, this experiment explores the effects of implementing FA practices in the classroom, in terms of students' text comprehension abilities. A two-group experimental design was carried out involving students from two first-year classes of a lower secondary school in Italy, who were randomly assigned to the treatment or control group. After administering a task on text comprehension abilities, the researcher developed 15 FA activities in the treatment group. At the end of the study, the students' abilities were measured again and the pre- post-test

* The overall structure of the contribution is the result of a shared reflection of the authors. Paragraphs 2, 3, 4, 5, 6, 7, 8, and 9 were written by E. Guasconi. Paragraphs 1 and 10 were written by I. Vannini.

ECPS Journal – 31/2025 - https://www.ledonline.it/ECPS-Journal/
Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

difference between the two groups was verified. Non-parametric tests were run that did not reveal statistically significant differences. Regardless, a trend shows a slightly higher increase in the achievement of the treatment group.

Keywords: Effective teaching; Formative assessment practices; Lower secondary school; Text comprehension abilities; Two-group experimental study.

1. Introduction

International debate underlines the strategic relevance of formative assessment (FA) within the teaching-learning processes. It is characterized as a process that takes place during the development of a teaching unit, and which allows teachers and students – thanks to formative feedback – to fill the learning gaps and «to adjust ongoing teaching and learning to improve students' achievement of intended instructional outcomes» (Bennett, 2011, p. 6).

Although a single definition of FA is not yet shared, many studies have highlighted the effectiveness (in terms of learning quality and equity) of specific classroom assessment practices linked to a formative function, i.e., an assessment that allows to collect, interpret and use evidence about students' achievement in order to make decisions to foster and improve their learning (Wiliam, 2011b; Black & Wiliam, 2018).

In the Italian debate, the hypothesis of the effectiveness of FA is very relevant, especially for those student skills with worrisome shortcomings, such as reading skills. In fact, the data on students' achievements show the need to improve students' reading skill levels in both primary and secondary schools; therefore, which strategies can foster students' achievements must be urgently investigated, and FA could be a relevant opportunity.

This purpose is also in line with the national public debate on assessment. Indeed, assessment practices have always been a critical element in the Italian school system (Vannini, 2012), despite the fact that some important regulations have sought to contrast a traditional and selective vision over the decades.

In this regard, educational research has not yet provided specific evidence on how FA works in Italian school contexts, and currently only some studies (none of them experimental) examine the use of FA practices in classrooms. For this reason, in this study we investigated – through an experimental design in two groups – the efficacy of FA starting from causal hypotheses already explored in other countries. In particular, the research

question is: how much can reading skills be improved in Italian middle schools through the use of FA in the classroom?

In the following sections, the variables examined are precisely defined. First and foremost, an operational definition is provided of the FA strategies considered as an independent variable.

2. Formative assessment practices to improve students' achievements

The roots of the interest in the relationship between formative assessment practices and students' achievements could be found in the early studies on Mastery Learning, the individualized instruction program theorized by Bloom. As the author reported (Bloom, 1971; Bloom *et al.*, 1971), its aim was to bring all students to master certain essential abilities using different methods and managing time according to their needs. FA had a critical role in pursuing this goal, since it allowed teachers to collect evidence on each student's learning to perceive any weaknesses and provide immediate support by planning appropriate correctives. Indeed, Guskey (2007, 2010) has stressed the importance of the cycle composed of «formative assessment – feedback – correctives or enrichment activities» within the program. Several studies from different countries have supported Bloom and colleagues' theory with empirical evidence. Later Black and Wiliam (1998a) used some of those results in a research review to enhance the claim regarding the efficacy of FA practices, and published the main findings of a meta-analysis reporting an effect size between 0.4 and 0.7 for those practices (Black & Wiliam, 1998b). They underlined the role of feedback in the FA process – which refers to an action that uses collected evidence on students' abilities to identify information about their strengths and weaknesses in order to suggest proper strategies for achieving the learning goals – following the trends of the scientific debate in those years, as several researchers were trying to understand which feedback feature could foster student motivation, interest and achievements between 1980 and 1990 (Deci & Ryan, 1982; Butler, 1987; Bangert-Drowns *et al.*, 1991; Kluger & DeNisi, 1996). The relevance of this element is confirmed by the fact that it's still at the center of the debate (Wisniewski *et al.*, 2020). According to Sadler's attention to student comprehension of the teacher-provided feedback (1989), several authors are now investigating how students can interpret and use it for their own improvement (Lipnevich *et al.*, 2016; Lipnevich & Panadero, 2021; Lui & Andrade, 2022).

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

After the 2000s, the search for an answer to the question regarding the effectiveness of FA practices has seen different contributions, including experimental and empirical studies as well as meta-analyses on «what works» in teaching. Concerning the former, some math and science research deserves mention, such as that carried out by Yin *et al.* (2008), Andersson & Palm (2017), and van den Berg *et al.* (2018), as well as studies on Keeping Learning on Track, a professional FA development program for teachers (Bennett, 2011). Only part of the results from these studies goes in the same direction, since some of them did not show any significant effect. Furthermore, they are concerned with FA practices in different subject areas. The same variability could be found in the second cited contributions on the efficacy issue: in Hattie's meta-analysis (2009), FA and formative feedback strategies register an effect size between 0.73 and 0.90, but this result is not consistent with that merged from Kingston and Nash's work (2011) which, starting from Black and Wiliam's claim (1998b), shows an effect size of FA practices of 0.25.

In considering these results, it is crucial to recognize the role of the context, with its peculiarities that could affect how practices are implemented in the classrooms and how students react to them. We must admit that the Italian school system sees a lack of a robust assessment culture where teachers collect evidence on students learning to make both their teaching and each student's learning «visible» in order to adjust activities. Moreover, assessment practices are still far from being systematic and based on a valid and reliable process, and teachers often make decisions about teaching that need to be better founded.

3. Which formative assessment practices? Defining strategies and assumptions

Bennett (2011) has underlined the need for a clear definition of FA (Cizek *et al.*, 2019): as a matter of fact, many authors refer to different strategies when citing it.

Wiliam and Thompson (2007) break the concept down into five different strategies by intersecting the three subjects mainly involved in a teaching-learning situation: teacher, student and peer, with the three questions representing the FA cycle (Hattie & Timperley, 2007). Within this framework, teachers share learning intentions with students and elicit evidence about their learning using different tasks and techniques. Then they provide feedback with suggestions for achieving the goals.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Later, Bennett (2011) used this construct to clarify which FA strategies are included in the Keeping Learning on Track program. In addition to what has been stated by the two authors, he stresses that the teacher must use the acquired information to adapt instruction, and that the students must act on the feedback. Further, van den Berg *et al.* (2018) underline the value of planning remedial activities after observing learners' understanding, whereas Allal and Lopez (2005) shift their attention from «what teachers can do» to «what students can do» to improve their learning. Finally, FA strategies that derive from what is called by Ruiz-Primo and Furtak (2006) «informal formative assessment» consider practices capable of enhancing interactions between teacher and students and between peers in order to co-construct learning.

Several issues arise from the study of FA practices. First, each construct must be considered part of an «assessment paradigm». Van der Klej *et al.* (2015, p. 237) explain that the expression refers to the aim and role assigned to the process. What is the purpose of assessment? The answer to this question leads to decisions about which methods would be used. Second, FA practices are rarely described in detail in research: clarifying FA strategies does not automatically lead to understanding of what teachers or researchers did in classrooms. Lastly, the relationship between disciplines and FA practices is now being investigated (Cizek *et al.*, 2019).

It must be stated that Dewey's theories on the democratic aim of education (1916/2018) form the foundation of the current research: the author explained that without education, individuals cannot ensure the intellectual and moral development they need to live and to improve society (Baldacci, 2017). Nevertheless, providing each student with the essential skills for participating in society is also recommended by the OECD and the Italian government (MIUR, 2012; 2018). Despite these recommendations, significant gaps exist between the results of advantaged and disadvantaged students in Italy (INVALSI, 2023; OECD, 2023a). Facing this discrepancy many years ago, in 1976 Vertecchi looked to Bloom's theories on the Mastery Learning program and to formative assessment, feedback, and correctives, as did Visalberghi (1955), Calonghi (1983), and after them Gattullo (1968), Domenici (2003), and as today do, among others, Benvenuto (2003, 2007), Agrusti (2021) and Vannini (2009, 2022). Within their studies, measurement is strongly emphasized as crucial for enacting a systematic and valid assessment process capable of catching and understanding students' real needs. If this is true for the «summative» assessment, it is also fundamental for the «formative» one, since valid and reliable data on each student's learning are resources for those who want to help them improve toward established goals when there is still time to do so.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Assuming this approach, FA is described in this project as *an ongoing process that starts from sharing learning objectives and success criteria with the students and sees the teacher gathering evidence on each student's learning to regulate the teaching-learning process. This happens primarily through feedback containing individualized strategies to close the possible gap between each student's level and that desired, as well as through the proposal of remedial or enrichment activities that meet students' needs. Secondarily, the improvement is possible by structuring moments where the students can reflect on their own and peers' learning*. In line with this definition, the construct is inspired by that created by Wiliam and Thompson (2007), but also contains references to the measurement and use of evidence collected by teachers in planning remedial activities (*Tab. 1*).

Table 1. – The construct of FA practices the research referred to (readapted from Wiliam & Thompson, 2007). Source: Guasconi, 2025.

Teacher	FA key questions	Student
Identify specific learning goals and assessment criteria and share them with students.	*Where is the learner going?*	Understand learning objectives and assessment criteria
Create and use valuable learning monitoring tools that allow to check every student's acquisition. Establish measurement criteria and implement adequate actions. Design opportunities for reflection on its own and peers' learning process (peer and self-assessment moments).	*Where is the learner right now?*	Engage in the tasks/tests proposed by the teacher. According to the criteria, reflect on your and your peers' learning tasks paying attention to the process and the applied strategies. Provide feedback to peers or/and find improvement strategies for them and for yourself.
Communicate individual/collective feedback (including individualized strategies to improve learning). Support each student in the use of suggested strategies. Based on collected evidence about each student's learning, plan remedial or enrichment activities.	*How to get there?*	Understand the teacher's feedback and reflect, together with the teacher, upon the distance between your assessment and the one made by the teacher. Catch the gap between your own level and the one represented by quality indicators. Use improvement strategies suggested by the teacher or found by yourself.

The framework has specific components related to self and peer assessment, but this aspect detects a secondary role among FA practices implemented in this experiment. The reasons lie in organizational issues: to see the effect of these specific practices, a more extended period than the one available would have been required since the students must know how to compare their and their peers' performance with that targeted, must be

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

able to analyze mistakes to find suggestions and must acquire metacognitive knowledge and abilities.

4. Formative assessment practices to enhance reading comprehension abilities

In classrooms, FA practices are related to a specific domain. This is also true for this study in which the practices are related to the linguistic area, and more specifically, to certain reading comprehension abilities.

We can state that comprehension is a complex process based on a two-way interaction between the reader and the text that leads the former to create a representation of the meanings of the text passage. It involves cognitive functions such as memory, attention and learning and contains different levels of complexity (De Beni & Pazzaglia, 1995; Cornoldi *et al.*, 2017). In this process inferences have a crucial role (Kintsch, 1988): they enable the integration of meanings and strengthen relevant connections between information in a text. They are also essential components of the reading literacy frameworks provided by the OECD (2023b), IEA (Mullis & Martin, 2019), and INVALSI (2018). Indeed, every institute locates the ability to make local and global coherence inferences within the process of «integration» of the meanings and information of the text (Cardarello & Bertolini, 2020). At a local level, it requires making what are known as «bridging inferences» (Oakhill & Cain, 2018) or «text-connecting inferences» (Cain & Oakhill, 1999), an integration of two or more near propositions also based on mapping related expressions, such as a pronoun with the name to which it refers. At a global level, integrating the text's meanings requires synthesizing and selecting relevant propositions (adding personal knowledge if necessary) in order to build a single coherent representation.

Oakhill and colleagues have confirmed the necessity of this kind of inference, since their research has come to the following conclusions (2015-2021), consistent with the ones reached in Italy by Lumbelli (2009).

- The ability to make inferences affects comprehension; *less skilled readers* struggle to recognize inappropriate inferences when they are put in a text.
- *Poor comprehenders* show difficulties understanding cohesive connections at a local level.

Since it would be impossible to consider all the reading comprehension abilities in the current study, the choice has fallen on those related to making local and global inferences precisely for their role in determining the understanding of a text. *Table 2* describes the deemed abilities and shows the corresponding learning objective for each of them.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Table 2. – Reading comprehension framework: selected abilities with corresponding learning objectives. Source: Guasconi, 2025.

Abilities	Learning objectives
Retrieve information that is explicitly stated in the text	Retrieve and locate specific information in the text.
Recognize concepts/meanings expressed in paraphrastic forms (direct inferences or first-level inferences)	Identify the required information, even when expressed in different ways, and employ simple reasoning processes to identify it.
Make lexical inferences	Understand the meaning of a word or expression from the context.
Make text-connecting inferences (coherence)	Establish the reference of an anaphors or an anaphoric chain.
	Discover the implicit connections between the information of the text mainly using text's clues.

The first ability in the table has been included because it is a prerequisite for identifying cues in a text that are helpful in integrating its meanings. Apart from this, the remaining abilities are related to making inferences, even if they have different levels of complexity and refer to specific tasks. Among the others, the recognition of paraphrastic forms in the text was the most difficult to define because it responds to the attempt to include a «simple» inferential process, which requires the reader to use cues to recognize that the concept expressed in the question has the same meaning as one existing in the text.

5. Aim of the study

The aim of the research was to explore the effect of the use of FA practices in Italian middle schools on the mentioned students' reading abilities. The study investigated whether using FA strategies in classrooms could enhance those abilities through an experiment with two randomized groups (Guasconi, 2025). It focused on lower secondary school to find strategies and practices able to improve students' comprehension abilities, which could eventually become the subjects of a structured pre-service and in-service training program. In fact, research shows that using FA practices in classrooms could be a valuable way to increase students' achievements, but their effectiveness must first be investigated in the specific Italian context.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Indeed, the national debate regarding which conditions allow FA practices to express their potential must urgently be pursued: a challenging goal that may not be satisfied with a single experiment, but which recalls several studies.

This project focused on the main issue regarding the effects of the use of FA practices in the classroom, and its specific questions related to the research design were:

- *RQ1 – Has the employment of FA practices with students of the experimental group (EG) generated a more extensive, positive, and statistically significant difference between the pre-test and post-test scores compared to the difference revealed for the control group (CG)?*
- *RQ2 – How much are FA practices relevant for improving students' specific reading abilities?*

6. Method

This section describes the characteristics of an experiment designed to validate a research hypothesis that assumed FA practices would positively affect students' reading comprehension abilities in Italy. The diagram in *Figure 1* summarizes the variables considered and the relationships among them.

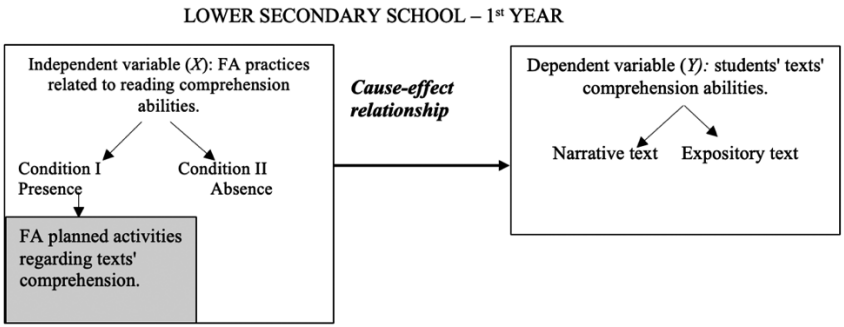


Figure 1. – Research hypothesis.
Source: Guasconi, 2025.

The independent variable has been identified in FA practices related to the reading comprehension domain, and it presented two modalities: presence or lack of planned FA activities. It was intended to affect students' reading comprehension abilities, a quantitative variable classified as dependent.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

6.1. *Participants*

The research was carried out in 2021-2022. It involved 47 students aged 10-11 who were attending the first year of a lower secondary school in the northwest part of Italy. Participants were randomly assigned to the experimental (N 25) and control group (N 22): the first was part of the treatment carried out by the researcher, whereas the other participated in teaching activities on the comprehension and interpretation of Greek myths developed by a teacher, a condition decided to meet the curriculum of the institute. The random assignment was accomplished by selecting the first or last half of the students' names from two separate classrooms in alphabetic order (the first letter of the surname) in which they were reported in each register and by randomly assigning the selected students to the experimental or control group.

Before forming the two groups, a series of meetings with the school's head teacher and his collaborators were held to share the research design and understand how to arrange the treatment to match and integrate the school's activities.

The University of Bologna Ethics Commission authorized the research in 2021, and the students' families were informed. The teachers and students were not specifically aware of the research hypothesis in order to avoid a reaction from them that could distort the results. The researcher let the students know that they were part of a university project aimed at improving teaching the Italian language. At the same time, the teachers acknowledged that activities would focus on reading comprehension abilities.

Students with disabilities and learning disorders were involved in the project and the treatment activities, but their data were excluded from the analyses. Data on students' abilities were anonymous: a code was randomly assigned to each student to match the pre-test and post-test data.

6.2. *Research design*

The experiment followed the design described by Campbell and Stanley (1963) as a «pre-test - post-test control group design» (p. 13), and its main phases are shown in *Figure 2*.

Before proceeding, some premises must be acknowledged.

After the definition of the theoretical hypothesis, the primary need was to narrow down the independent variable (Lumbelli, 1984; 2006), searching for elements and actions that compose it. That was precisely the aim of the *exploratory phase* of the project, carried out between November 2021 and February 2022 and divided into two parts.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

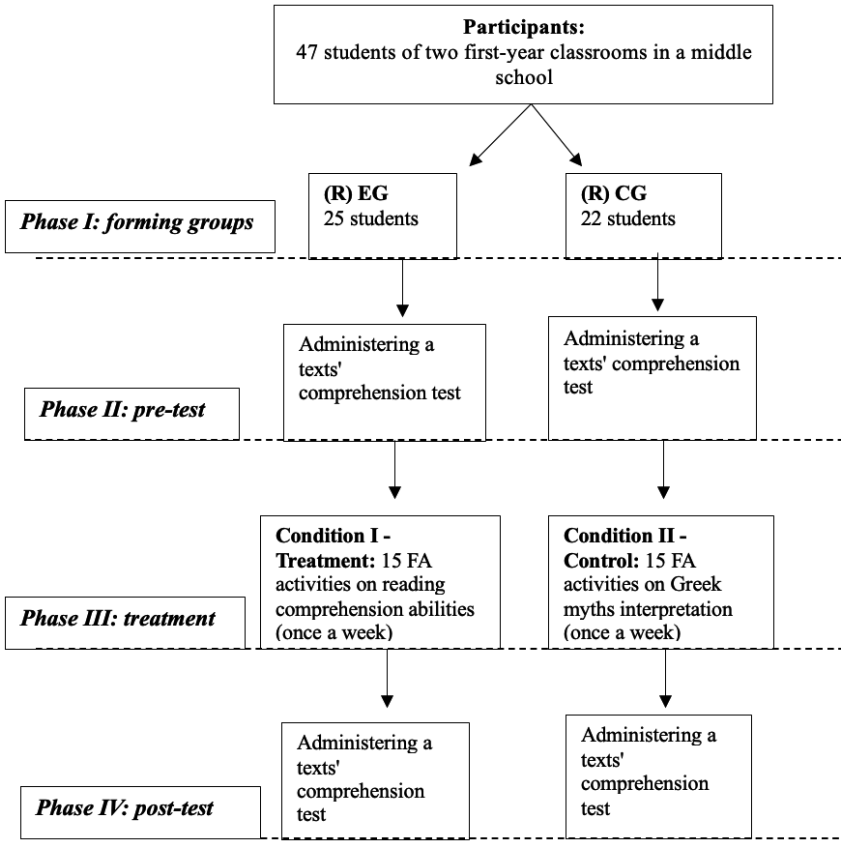


Figure 2. – Research design. EG = Experimental Group; CG = Control Group; R = Randomized.

- The first was conducted in a middle school in the same region as that involved in the experiment. After identifying 80 practices (see the example of the specific practices in *Tab. 3*) from the strategies included in the FA construct, the researcher implemented them in 14 short FA activities related to some of the selected comprehension abilities with first-year students between December 2020 and May 2021. A questionnaire was administered to seize students' perceptions and opinions about the techniques and materials used, and a research journal was kept to note of any difficulties.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Table 3. – FA practices related to the strategy identify specific learning goals and assessment criteria and share them with students.

FA strategy	Indicators of FA practices – micro variables (behaviors)
C Identify specific learning goals and assessment criteria and share them with students	C1 The researcher identifies the macro-skills, the specific learning objectives (the selected reading comprehension abilities), and the assessment criteria.
	C2 The researcher shares with students the macro-skills (learning goals) on which the FA activities will focus paying attention to using a language that is understandable for students of that age.
	C3 The researcher uses multiple languages for the presentation and explanation of macro-skills (*verbal and non-verbal; for example, the researcher can use iconic mediators*)
	C4 At the beginning of each FA activity, the researcher illustrates the specific learning objective (object of assessment) to students using comprehensible language (*what should they know and/or what should the students be able to do?)* and emphasizes the connections with the macro-skills.
	C5 The researcher uses multiple languages for the explication of learning objectives (*verbal and non-verbal; for example, the researcher can use iconic mediators*)

- The second was carried out separately in two first-year classrooms of the same school involved in the experiment, between October 2021 and February 2022. Two quasi-experimental plans were carried out: they employed quantitative analyses of the difference between the experimental and control groups and qualitative analyses of the experimenter's journal.

At the end of the exploratory phase, 15 FA activities related to text comprehension abilities were planned in detail to shape the treatment of the experiment. The research design followed the subsequent phases (*Fig. 2*).

- *Group composition*: participants were randomly assigned to the experimental and the control group. The small number of students per group must be highlighted; if the excluded students are subtracted from the total, the number of students is 18 for the experimental group and 16 for the control group. Moreover, descriptive analyses have confirmed they have different characteristics (see § 8).
- *Pre-test*: administration of a reading comprehension test composed of 14 items: nine referred to a narrative text and five to an expository text. Following the planned conditions, the test was simultaneously administered to the control and experimental groups in the second week of February 2022. All the questions (except one) were taken from INVALSI tests from the year 2013-2014; they were multiple-choice or requested a unique answer (only one required an open response).
- *Creation of the two conditions of the independent variable*:
  - Condition I involved EG students and saw the implementation of the 15 planned FA activities. Each lasted two hours (between 2 and 4 pm)

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

and was carried out once a week from the 15th of February 2022 to the 18th of May 2022. The activities were carried out by the experimenter, who in this research was the same person as the researcher (see § 7 and *Tab. A1* in Appendix A for a description of the intervention).
  - Condition II consisted of 15 didactic activities (carried out simultaneously and with the same frequency as condition I) regarding the reading and interpretation of Greek myths. No assessment moments were developed, as the teacher confirmed in an interview at the end of the activities (see *Tab. A2* in Appendix A).
- *Post-test*: this phase included the administration of the same test used as a pre-test (the time interval was three months), called «post-test B», and of another parallel test after that, called «post-test A». Items of post-test A had similar levels of difficulty, the same number and order as those of post-test B, but the texts were different and belonged to INVALSI tests from a different year (2014-2015) (see Appendix B, *Tabs. B1* and *B2* for details regarding the two tests). Employing two similar tests allowed the researchers to further reflect on the main results obtained with the test-retest procedure. Following the planned conditions, the tests were simultaneously administered to the control and experimental groups at the end of May 2022. However, it must be said that there was no real parallelism between the two reading tasks because the correlation between students' scores obtained with the post-test B and the post-test A hadn't been monitored before the administration, but only at the end of the experiment for time reasons.

The randomization of the groups is one of the major strengths of the project, as well as the treatment period (three months) and the use of two tests in the post-test phase. In contrast, the small number of students, together with some features of condition II (not optimal for controlling the effect of the independent variable) are its main limits. Furthermore, it's important to mention that the FA activities fully corresponded to the strategies listed in the *Table 1*, which represents the essential elements of an assessment process as described by Gattullo (1968): the definition and the sharing of the assessment objects (learning objectives) and criteria; the measurement phase consisted of planning and using learning tasks that are valid for what concern the assessment objects and reliable, and lastly, the interpretation of the measurements and the generation of formative feedback (only during an FA activity). This means that FA activities implemented in the treatment were not included in a specific structured teaching unit. Although the inclusion of these activities in an instructional cycle would add more value to the research (Wiliam, 2011a), unfortunately, it was not possible at the time of the project implementation.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

6.3. *Analysis*

The development of experimental or quasi-experimental studies in education is often done in settings with small sample sizes and/or a distribution of dependent variable measures that is different from normal. In such cases, analyses using non-parametric tests are more appropriate (Domenici, 2006) since the requirements for making inferences on the population are not met. For this reason and due to the small number of students, non-parametric tests were employed (Rajaretnam, 2016). After some descriptive analyses, the Mann-Whitney U test was run considering the difference between the two (independent) groups (EG and CG) in the pre-test - post-test discrepancy. How much variability in the difference between the pre-test and the post-test is due to the presence or absence of FA practices? How much is this variance statistically significant? The test answered these questions.

Moreover, effect size (ES), an index for seizing the practical significance not based on statistical significance, was computed.

All the analyses have taken into consideration both the difference between pre-test and post-test B and between pre-test and post-test A. Here the results examine the first difference or the test-retest procedure, but some mention will also be made on the analyses related to the other.

7. The intervention

The planned FA moments could be divided into three parts: the presentation of the entire set of activities and FA strategies (named «cycle» to underline its cyclical nature) (1 moment); the development of the activities, and the conclusion of the treatment (14 moments).

During the presentation of the activities, the experimenter used different texts and posed questions to elicit dialogue from students concerning how much adequate comprehension is essential in daily life and the consequences of improperly interpreting a text. Then she used the metaphor of a puzzle, with the support of a billboard prepared for this purpose, to share targeted learning objectives with students and presented the cycle of FA activities as follows.

- «What are we learning to do?» – Sharing learning objectives and assessment criteria.
- «What is our current level?» – Formative assessment tasks.
- «What can we learn from monitoring?» – Formative feedback.
- «Trying to improve!» – Using suggestions and strategies to improve.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

During the first two moments, the students also personalized a pre-structured learning journal to document their process (see *Fig. A1* in Appendix A). After that, each activity was divided into the following phases: a brief recall of previous arguments, the presentation of a timetable regarding tasks scheduled for the day; and the actual activities.

To share the learning objectives, the experimenter used many examples of texts (often read aloud and supported by images and videos) accompanied by questions: the students were invited to guess the answer and which ability was needed in order to find a response. The objectives were also translated in different ways (also using role-playing), read, and explained. The reflection on assessment criteria instead used exemplars and rubrics: in the first case, students were encouraged to say why a task was «well or poorly done», while in the second one, they tried to use a rubric to provide some fictional peers with suggestions for improvement, according to the task they were analyzing. The FA tasks were introduced by recalling their aim to improve students' learning, the specific goal, and by explaining the instructions. Most of them included a narrative and an expository text with multiple-choice questions and lasted 30 minutes. Often, they also contained a semi-structured item (Domenici, 2005) that allowed students to explain why they gave the answer in order to understand their reasoning and use it to structure appropriate feedback.

Before every FA activity, the experimenter expressed her trust in the students' abilities and underlined the value in the role of their engagement. Texts selected from experts and from INVALSI were taken to create formative assessment tasks.

Once the FA tasks were administered, the students' answers were registered and analyzed using a spreadsheet. It has been structured in such a way as to be able to code errors, note the improvement strategies, calculate measures of central tendency and variability, and calculate item analysis.

In two FA cycles (regarding the following abilities: «identify information that is explicitly stated in the text» and «make text-connective inferences – identify the reference of anaphors and anaphoric chains») individual written feedback was given to the students, which was also discussed later with each of them. In the remaining cycles, collective oral and collective written feedback (then reported by students in their learning journal) were provided.

During each feedback moment, the researcher used three symbols to better represent the concepts of strengths (a star), weaknesses (a cloud), and strategies for improvement (a balloon). In providing individual feedback, the researcher read the comments together with each student, paying

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

particular attention to enhance reflection and to deepen the student's reasoning process. Then the researcher checked the students' understanding of the feedback received: if they had not understood how to use the recommended strategies, the researcher adopted *modeling* techniques.

In all the FA cycles (except one), the exercises proposed after giving the feedback asked students to re-answer the questions they had originally answered incorrectly. Indeed, the correct answer was not revealed, so the researcher asked them to reconsider some questions using the recommended strategies.

Table A1 in Appendix A shows some examples of macro-activities carried out in the intervention.

8. Results

After the exclusion of students with learning disorders and absent during the pre-test or the post-test phase and in order to explore the discrepancy between the two groups of the pre-test - post-test B difference, the 34 valid cases were divided as follows: N(EG) = 18; N(CG) = 16.

The students of the EG and CG showed different situations in the pre-test.

As *Table 4* reveals, in the pre-test the students from the EG achieved a mean score of 10.44 (the maximum score was 14) with a standard deviation value of 3.03. In contrast, the CG group students reached an average score that was slightly lower, 10.25, with a lower standard deviation value (2.27). Looking at the mode, it is also noted that this value is very different for the pre-test (14 for the EG and 7 for the CG) in the two groups.

Table 4. – The control (a) and experimental group's (b) descriptive analyses.

(a)
CG – Control Group (N 16)

	Pre-test	Post-test B
Mean	10.25	11.25
Median	10.50	11.50
Mode	7	12
St. Deviation	2.27	1.84
Skewness	-0.08	-0.20
Kurtosis	-1.07	-0.90

(b)
EG – Experimental Group (N 18)

	Pre-test	Post-test B
Mean	10.44	11.72
Median	11	12
Mode	14	13
St. Deviation	3.03	2.05
Skewness	-0.42	-0.68
Kurtosis	-1.07	0.61

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

The post-test B data show a slight increase for both groups: the mean score for the EG students in the test is 11.72, and for the CG is 11.25, showing a higher value for the former; the variability declined in both groups: in post-test B the registered standard deviation is 2.05 for the EG and 1.84 for the CG.

Examining the distribution of the students' scores in the pre-test is also interesting: six of them in the EG showed difficulties with the test (in other words, they scored less than 10 points), reaching a mean of 6.83; they then increased their results in post-test B, as their mean score rose (9.83). The same number of students had trouble with the pre-test in the CG, registering a mean score of 7.83, but in post-test B they increased their results less than the other group's students (9.66).

The analyses concerning post-test A almost confirmed the same path: students belonging to the EG performed slightly better than the others, showing higher values for the mean score in the post-test than the CG, even though this test was more difficult for all, and they show a bigger increase. As for the difference between post-test B and the pre-test, EG also reduced the variability in results more than the CG which, on the contrary, increased it.

Following these descriptive analyses, answers to the research questions are reported below.

RQ1 – Has the employment of FA practices with students of the EG generated a statistically significant difference between the pre-test and post-test scores compared to the difference revealed for the CG?

The pre-test - post-test B difference reported an average of +1.28 for students of the EG and of +1 for those of the CG.

The Mann-Whitney U test confirmed that this difference was not affected by belonging to the EG or CG in a statistically significant way. Almost the same pattern emerged from the difference between pre-test and post-test A (in order to calculate the difference between pre-test and post-test A, the scores have been transformed into standardized scores), as the EG students registered a slightly higher increase than the CG, but which was also not statistically significant.

However, it is also interesting to explore the results of the Mann-Whitney U test of the variance between the two groups (belonging to the EG and the CG) of low-achieving students in the pre-test - post-test B difference. Indeed, students who had more trouble during the pre-test and reached the lowest scores (between 5 and 10) belonging to the EG saw relevant growth (+3 points) on average in post-test B, higher than that reached by students in the CG (+1.83 points). In this case, the difference between the two groups is nearly statistically significant ($p = 0.067$).

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

RQ2 – How much are FA practices relevant for improving students' specific reading abilities?

The analysis of the practical significance provides a measure of the effectiveness of FA intervention in the specific context. It was carried out considering both the differences between pre-test - post-test B and the pre-test - post-test A (T scores). The comparison between pre-test and post-test B showed a minimal effect (0.18) (*Tab. 5*). This result is also confirmed by the effect size (ES) value that refers to the difference between pre-test and post-test A, which is slightly higher (0.23) but still remain low to denote an effect of the FA activities.

Table 5. – (a) Effect size's values of the difference between the post-test B and pre-test. (b) Effect size's values of the difference between standardized scores of post-test A and pre-test.

(a)
Post-test A - Pre-test difference

		Confidence interval (95%)	
		Inf.	Sup.
Cohen's d	0.19	-0.49	0.86
Hedges' g	0.18	-0.48	0.84
Glass	0.25	-0.43	0.93

(b)
Post-test B - Pre-test difference

		Confidence interval (95%)	
		Inf.	Sup.
Cohen's d	0.23	-0.46	0.92
Hedges' g	0.23	-0.44	0.89
Glass	0.24	-0.45	0.93

9. Discussion

It is clear that the results suggest one main answer to all the questions: the FA activities carried out during the treatment phase did not significantly change students' reading abilities; for this purpose, we maintained the null hypothesis. However, if this is the big picture, we must also say it has many nuances.

Above all, data reveal a higher increase, albeit small, for students in the EG compared to the one which emerged from the other group (CG). This is also true for the analyses carried out on the pre-test - post-test A difference and confirmed a tendency already noted in the results of the two quasi-experimental plans where the treatment groups from each classroom had achieved a higher growth in post-test B, compared to that reached by the students of the control groups. Moreover, it is crucial to note that in the experiment, the EG students who raised their results in post-test B were those who struggled more at the beginning (a growth of 3 points), and to observe that this improvement was not so strong in the CG (1.83 points).

The same can be said for the decline of the standard deviation in the EG, especially if the theoretical assumption of the democratic aim of education is considered.

Secondly, observing the groups' characteristics enhances some issues. It is difficult to carry out statistical analyses with such a restricted number of cases. In addition, at the beginning of the experiment the EG students included seven who scored 12-14 points on the pre-test. Therefore, they probably would need more enhancement and more challenging activities. Assuming this perspective, it would be plausible to see a slight improvement since seven already had good reading abilities.

The ES values confirm the inference regarding the fact that the FA activities have slightly improved the reading abilities of the students of the EG. According to Hattie (2009), the ES must be greater than 0.40 (the «hinge point») to signal a visible efficacy. Anyway, other authors have classified 0.20 as a «small effect» (Pellegrini *et al.*, 2018). The scarce control of what has been called «condition II» (the alternative one) of the independent variable, together with the isolation of the FA practices, which were not embedded within a teaching-leaning unit, is part of the reasons for the registered magnitude of the FA efficacy.

Many intervening variables occurred during the experiment as well, such as the Covid-19 emergency that caused the absence of some students of the EG from a few of the activities of the intervention.

10. Conclusion

The article presented an experimental design to verify the influence of FA strategies on reading comprehension skills in Italian lower secondary schools. In the framework of the international debate, the construct of FA has been defined as a set of specific practices in the classroom: based on Wiliam and Thompson's construct (2007), the experimenter carried out practices to clarify the learning objectives and criteria for assessment; monitor student achievement; communicate formative feedback and support students in the use of the improvement strategies.

Reading comprehension skills were clearly defined as variables and indicators (identify information, recognize paraphrastic forms, make lexical and text-connecting inferences), measured at input and output through validated tests.

The results of the experimentation highlight the strengths and weaknesses of the design.

The structural difficulties of Italian schools in carrying out experiments led to having a very small number of subjects in the CG and EG. This severely limited the possibilities for interpreting the results obtained. On the other hand, the data analysis (also through effect-size measurements) shows some interesting improvements in the EG, where internal variability also decreases. A more marked increase in results among the EG students who had more trouble with the pre-test should also be noted. All this contributes to supporting the hypothesis of the effectiveness of FA in improving teaching-learning results in the field of reading comprehension.

Despite the critical points of the experiment, the results reveal some interesting interpretations to continue exploring in the future, replicating the experimental design on larger groups of students, in order to obtain results with a good degree of external validity. This is very important, especially in the Italian context where experimental research, with solid designs and valid measurements, is lacking.

References

Allal, L., & Lopez, L.M. (2005). Formative assessment of learning: A review of publications in French. In *Formative assessment: improving learning in secondary classrooms* (pp. 241-264). Paris: OECD Publishing.

Agrusti, G. (2021). Le prove di verifica. In E. Nigris & G. Agrusti (a cura di), *Valutare per apprendere* (pp. 78-88). Milano - Torino: Pearson.

Andersson, C., & Palm, T. (2017). The impact of formative assessment on student achievement: A study of the effects of changes to classroom practice after a comprehensive professional development programme. *Learning and Instruction*, *49*, 92-102.

Baldacci, M. (2017). Democrazia e educazione. Una prospettiva per i nostri tempi. In M. Fiorucci & G. Lopez (a cura di), *Jhon Dewey e la pedagogia democratica del '900* (pp. 21-38). Roma: RomaTrE-Press.

Bangert-Drowns, R.L., Kulik, J.A., & Kulik, C.L.C. (1991). Effects of frequent classroom testing. *Journal of Educational Research*, *85*(2), 89-99.

Bennett, R.E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, *18*(1), 5-25.

Benvenuto, G. (2003). *Mettere i voti a scuola. Introduzione alla docimologia*. Roma: Carocci.

Benvenuto, G. (2007). *Gli esami sono maturi? Critiche e prospettive docimologiche per gli esami di stato*. Roma: Anicia.

Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7-74.

Black, P., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *80*(2), 139-148.

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, *25*(6), 551-575.

Bloom, B.S. (1971). Mastery learning. In J.H. Block (Ed.), *Mastery learning: Theory and practice* (pp. 46-63). New York: Holt, Rinehart and Winston.

Bloom, B.S., Hastings, T., & Madaus, G.F. (1971). *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill.

Butler, R. (1987). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. *Journal of Educational Psychology*, *79*(4), 474-482.

Cain, K., & Oakhill, J.V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing*, *11*(5), 489-503.

Calonghi, L. (1983). *Valutare. Risultati docimologici e indicazioni per la scheda*. Milano: Istituto Geografico De Agostini.

Campbell, D.T., & Stanley, J.C. (1963). *Experimental and quasi-experimental designs for research*. Boston, MA: Houghton Mifflin.

Cardarello, R., & Bertolini, C. (2020). *Didattiche della comprensione del testo. Metodi e strumenti per la scuola primaria*. Roma: Carocci.

Cizek, G.J., Andrade, H.L., & Bennett, R.E. (2019). Formative assessment: History, definition and progress. In H.L. Andrade, R.E. Bennett, & G.J. Cizek (Eds.), *Handbook of formative assessment in the disciplines* (pp. 3-19). New York: Routledge.

Cornoldi, C., Colpo, G., & Carretti, B. (2017). *Kit scuola secondaria di I grado*. Milano: GiuntiEdu.

De Beni, R., & Pazzaglia, F. (1995). *La comprensione del testo. Modelli teorici e programmi di intervento*. Milano: Utet Libreria.

Deci, E.L., & Ryan, R.M. (1982). Curiosity and self-directed learning: The role of motivation in education. In L.G. Katz (Ed.), *Current topics in early childhood education* (pp. 71-86). New York: Ablex.

Dewey, J. (2018). *Democrazia ed educazione. Un'introduzione alla filosofia dell'educazione* (ed. orig. *Democracy and education*, 1916). Trad. it. di G. Spadafora. Roma: Anicia.

Domenici, G. (2003). *Manuale della valutazione scolastica*. Lecce: Laterza.

Domenici, G. (a cura di). (2005). *Le prove semistrutturate di verifica degli apprendimenti*. Torino: Utet.

Domenici, G. (2006). *Metodologia della ricerca educativa*. Roma: Monolite.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Gattullo, M. (1968). *Didattica e docimologia. Misurazione e valutazione nella scuola*. Roma: Armando Editore.

Guasconi, E. (2025). *Prassi di formative assessment e comprensione dei testi*. Milano: FrancoAngeli.

Guskey, T.R. (2007). Closing achievement gaps: Revisiting Benjamin S. Bloom's «Learning for Mastery». *Journal of Advanced Academics*, *19*(1), 8-31.

Guskey, T.R. (2010). Formative assessment: The contributions of Benjamin S. Bloom. In H.L. Andrade & G.J. Cizek (Eds.), *Handbook of formative assessment* (pp. 106-124). London - New York: Routledge.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievements*. London - New York: Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81-112.

INVALSI – Istituto nazionale per la valutazione del sistema educativo di istruzione e di formazione (2018). *Quadro di riferimento delle prove INVALSI di Italiano*.
https://invalsiareaprove.cineca.it/docs/file/QdR_ITALIANO.pdf

INVALSI – Istituto nazionale per la valutazione del sistema educativo di istruzione e di formazione (2023). *Rapporto INVALSI 2023*.
https://invalsiareaprove.cineca.it/docs/2023/Rilevazioni_Nazionali/Rapporto/Rapporto%20Prove%20INVALSI%202023.pdf

Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues & Practice*, *30*(4), 28-37.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163-182.

Kluger, A.N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254-284.

Lipnevich, A.A., Berg, D.A.G., & Smith, J.K. (2016). Toward a model of student response to feedback. In G.T. Brown & L.R. Harris (Eds.), *Handbook of human and social condition in assessment* (pp. 169-185). London - New York: Routledge.

Lipnevich, A.A., & Panadero, E. (2021). A review of feedback models and theories: Descriptions, definitions, and conclusions. *Frontiers in Education*, *6*, 1-29.

Lui, A.M., & Andrade, H.L. (2022). The next black box of formative assessment: A model of the internal mechanisms of feedback processing. *Frontiers in Education*, *7*, 1-14.

Lumbelli, L. (1984). Qualità e quantità nella ricerca empirica in pedagogia. In E. Becchi & B. Vertecchi (a cura di), *Manuale critico della sperimentazione e della ricerca educativa* (pp. 101-133). Milano: FrancoAngeli.

Lumbelli, L. (2006). Costruzione dell'ipotesi e astrazione nella pedagogia sperimentale. In A. Bondioli (a cura di), *Fare ricerca in pedagogia. Saggi per Eagle Becchi* (pp. 25-60). Milano: FrancoAngeli.

Lumbelli, L. (2009). *La comprensione del testo come problema. Il punto di vista cognitivo*. Roma: Editori Laterza & Figli Spa (Kindle Edition).

MIUR – Ministero dell'Istruzione, dell'Università e della Ricerca (2012). *Indicazioni Nazionali per il curricolo della scuola dell'infanzia e del primo ciclo d'istruzione*.
https://www.mim.gov.it/documents/20182/51310/DM+254_2012.pdf

MIUR – Ministero dell'Istruzione, dell'Università e della Ricerca (2018). *Indicazioni Nazionali e nuovi scenari*.
https://www.miur.gov.it/documents/20182/0/Indicazioni+nazionali+e+nuovi+scenari/

Mullis, I.V.S., & Martin, M.O. (Eds.). (2019). *PIRLS 2021 Assessment frameworks*, Boston College. Chestnut Hill, MA: TIMSS & PIRLS International Study Center.
https://timssandpirls.bc.edu/pirls2021/frameworks/

Oakhill, J., & Cain, K. (2018). Children's problems with inference making: Causes and consequences. *Bulletin of Educational Psychology*, *49*(4), 683-699.

Oakhill, J., Cain. K, & Elbro, C. (2021). *La comprensione del testo. Dalla ricerca alla pratica* (ed. orig. *Understanding and teaching reading comprehension: A handook*, 2015). Trad. it. di G.M. Cataldo. Roma: Carocci.

OECD – Organization for Economic Cooperation and Development (2023a). *PISA 2022 results*, Vol. 1: *The state of learning and equity in education*. Paris: OECD Publishing.
https://doi.org/10.1787/53f23881-en

OECD – Organization for Economic Cooperation and Development (2023b). *PISA 2022 assessment and analytical framework*. Paris: OECD Publishing.
https://doi.org/10.1787/dfe0bf9c-en

Pellegrini, M., Vivanet, G., & Trinchero, R. (2018). Gli indici di effect size nella ricerca educativa. Analisi comparativa e significatività pratica. *Journal of Educational, Cultural and Psychological Studies (ECPS)*, *18*, 275-309.

Rajaretnam, T. (2016). *Statistics for social sciences*. New Delhi: Sage.

Ruiz-Primo, M.A., & Furtak, E.M. (2006). Informal formative assessment and scientific inquiry: Exploring teachers' practices and student learning. *Educational Assessment*, *11* (3-4), 205-235.

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

van den Berg, M., Bosker, R.J., & Suhre, C.J.M. (2018). Testing the effectiveness of classroom formative assessment in Dutch primary mathematics education. *School Effectiveness and School Improvement*, *29*(3), 339-361.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

van der Kleij, F.M., Vermeulen, J.A., Schildkamp, K., & Eggen, T.J.H.M. (2015). Integrating data-based decision making, assessment for learning and diagnostic testing in formative assessment. *Assessment in Education: Principles, Policy & Practice*, *22*(3), 324-343.

Vannini, I. (2009). *La qualità nella didattica. Metodologie e strumenti di progettazione e valutazione*. Trento: Erickson.

Vannini, I. (2012). *Come cambia la cultura degli insegnanti. Metodi per la ricerca empirica in educazione*. Milano: FrancoAngeli.

Vannini, I. (2022). *Una valutazione che analizza e ricostruisce. Rileggere oggi «Didattica e docimologia» in dialogo con Mario Gattullo*. Roma: Armando Editore.

Vertecchi, B. (1976). *Valutazione formativa*. Torino: Loescher.

Visalberghi, A. (1955). *Misurazione e valutazione nel processo educativo*. Roma: Edizioni di Comunità.

Wiliam, D. (2011a). *Embedded formative assessment*. Bloomington, IN: Solution Tree Press.

Wiliam, D. (2011b). What is assessment for learning? *Studies in Educational Evaluation*, *37*(1), 3-14.

Wiliam, D., & Thompson, M. (2007). Integrating assessment with learning: What will it take to make it work? In C.A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53-82). London - New York: Routledge.

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology*, *10*, 1-14.

Yin, Y., Shavelson, R.J., Ayala, C.C., Ruiz-Primo, M.A., Brandon, P.R., Furtak, E.M., Tomita, M.K., & Young, D.B. (2008). On the impact of formative assessment on student motivation, achievement, and conceptual change. *Applied Measurement in Education*, *21*(4), 335-359.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

APPENDIX A

Table A1. – The description of five activities of the intervention carried out in the EG by the researcher.

Moment (from 2 pm to 4 pm each)	Main activity	Detailed activities
1	*Presentation of the entire set of activities*	The experimenter and students' presentation. Reading of textual passages from the novel entitled «The Lord of the Rings» (Tolkien) followed by a collective reflection upon the relevance of a correct understanding. Activities structure's explanation with a description of the functions of the student's learning journal (*Fig. A1*).
2	*Presentation of the intervention FA cycle 1* *Learning objective: locate specific information in the text.*	Introduction of the organization and phases of a formative assessment cycle. Activities regarding the billboard entitled «The puzzle of the comprehension skill» including discussions on examples linked to the different objectives composing the skill (or, in other terms, the whole puzzle). Personalization by each student of the individual journal. The first learning objective's presentation and the sharing of the assessment criteria related to it (indicators of a quality performance).
3	*FA cycle 1* *Learning objective: locate specific information in the text.*	Introduction of the intervention's organization (through the «classroom daily schedule»). Random assignment of tasks to students (tidy classroom, distribute materials, write on the billboard …). A brief recall of the activities carried out during the previous moment (2) using the FA cycle already presented. Students write learning objectives and assessment criteria in their learning journals and report them on the billboard (translating them with their words). Presenting and performing the 1st formative assessment test (fast FA, since it comprises two questions). Based on students' answers, the distribution of individual written feedback (already prepared).

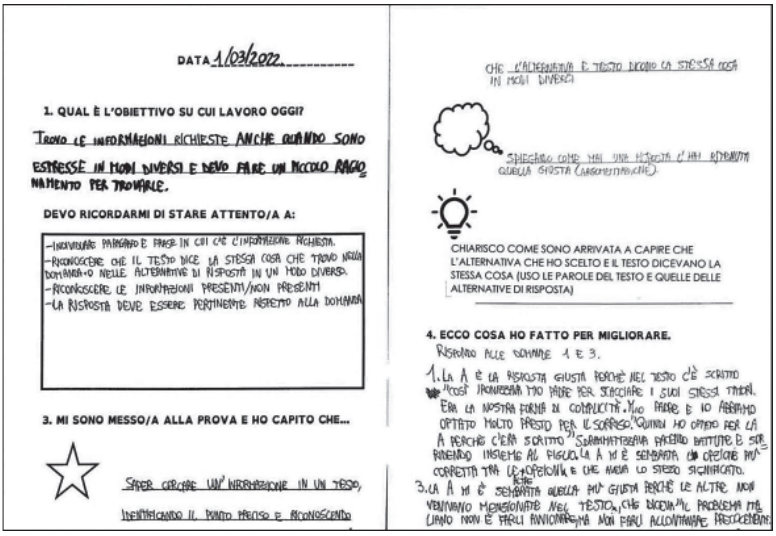
Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Moment (from 2 pm to 4 pm each)	Main activity	Detailed activities
4	*FA cycle 1/2* *Learning objectives: locate specific information in the text/ identify the necessary information, even when expressed in different ways, and use short reasoning to identify it.*	Introduction of the intervention's organization (through the «classroom daily schedule»). Random assignment of tasks to students (tidy classroom, distribute materials, write on the billboard …). A brief recall of the activities carried out during the previous moment (3) using the FA cycle. Students write the received feedback in their learning journals. Students make individualized exercises following the recommended strategies included in the feedback. The introduction of the new learning goal through reading aloud and discussing examples. Activities to focus students, who work in different groups, on the assessment criteria related to the new objective. In particular, the structuring of the roles within each group.
5	*FA cycle 2* *Learning objective: identify the necessary information, even when they are expressed in different ways.*	Introduction of the intervention's organization (through the «classroom daily schedule»). Random assignment of tasks to students (tidy classroom, distribute materials, write on the billboard …). A brief recall of the activities carried out during the previous moment (4) using the FA cycle. Group activity for commenting on a video showing two different comprehension task performances (second learning objective) explaining the reasoning processes in order to find hints regarding the assessment criteria. This activity is followed by a collective discussion. Students write the new learning objective and the assessment criteria in their learning journals and on the billboard. Introduction and administration of the 2nd formative assessment task.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Table A2. – The description of the activities carried out in the CG by a teacher belonging to the school as they were described by the teacher himself during an interview. It should be noted that the teaching blocks 2, and 3, outlined below, were repeated for the reading of three myths.

Teaching activities' blocks	Main activities	Detailed activities
1 (3/4 meetings, lasting from 2 to 4 pm each)	*Teachers' and students' presentation.*	Teachers' and students' presentation through playful activities.
	Activities to reflect and share personal interests, strengths and limits.	Drawings activities, sharing one's own story …
2 (1/2 meetings, lasting from 2 to 4 pm each)	*Collective reading of the 1st myth from a book.*	Students were invited to read a myth aloud; during the reading, they can write down words, phrases, and ideas that the story evoked in them, as well as draw places and characters. At the end of the collective reading, an oral summary of the myth is proposed, along with sharing the material produced by the students.
3 (2/3 meetings, lasting from 2 to 4 pm each)	*Audiovisual production (to be converted into a digital file) of the read the myth.*	Students were invited to produce an audiovisual presentation (to be converted into a digital file) that tells the story of the read myth in groups established by the teacher. To achieve this, they could use online software. The roles assigned to the students in the groups were not predetermined but were decided upon by each group/pair of students. The teacher constantly monitored the students' work during this activity by going around the groups.


DATA 1/03/2022

1. QUAL È L'OBIETTIVO SU CUI LAVORO OGGI?

TROVO LE INFORMAZIONI RICHIESTE ANCHE QUANDO SONO ESPRESSE IN MODI DIVERSI E DEVO FARE UN PICCOLO RAGIONAMENTO PER TROVARLE.

DEVO RICORDARMI DI STARE ATTENTO/A A:

- INDIVIDUARE PARLANDO E FARE IN CUI C'È L'INFORMAZIONE RICHIESTA.
- RICONOSCERE CHE IL TESTO DICE LA STESSA COSA CHE TROVO NELLA DOMANDA O NELLE ALTERNATIVE DI RISPOSTA IN UN MODO DIVERSO.
- RICONOSCERE LE INFORMAZIONI PRESENTI/NON PRESENTI
- LA RISPOSTA DEVE ESSERE PERTINENTE RISPETTO ALLA DOMANDA

3. MI SONO MESSO/A ALLA PROVA E HO CAPITO CHE...

SAPER CERCARE UN'INFORMAZIONE IN UN TESTO, IDENTIFICANDO IL PUNTO PRECISO E RICONOSCENDO

CHE L'ALTERNATIVA E TESTO DICONO LA STESSA COSA IN MODI DIVERSI

SPIEGARLO COME MAI UNA RISPOSTA C'HAI RIPENSATO QUELLA GIUSTA (CARATTERISTICHE)

CHIARISCO COME SONO ARRIVATA A CAPIRE CHE L'ALTERNATIVA CHE HO SCELTO E IL TESTO DICEVANO LA STESSA COSA (USO LE PAROLE DEL TESTO E QUELLE DELLE ALTERNATIVE DI RISPOSTA)

4. ECCO COSA HO FATTO PER MIGLIORARE.

RISPONDO ALLE DOMANDE 1 E 3.

1. LA A È LA RISPOSTA GIUSTA PERCHÉ NEL TESTO C'È SCRITTO COSÌ TRANNE MIO PADRE PER STACCIARE I SUOI STESSI PIANI. ERA LA NOSTRA PRIMA IN COMPAGNIA MIO PADRE E IO AVEVAMO OTTENUTO MOLTO PRESTO PER IL SOFFERTO QUINDI HO OPTATO PER LA A PERCHÉ C'ERA SCRITTO SOMMATTERANNA FACENDO BATTUTE E SORRIDENDO INSIEME AL FIGLIO LA A MI È SEMBRATA LA OPZIONE PIÙ CORRETTA TRA LE OPZIONI E CHE MANDA LO STESSO SIGNIFICATO.

3. LA A MI È SEMBRATA QUELLA PIÙ GIUSTA PERCHÉ LE ALTRE NON VENIVANO MENZIONATE NEL TESTO, CHE BLOCCAVA IL PROBLEMA MA LUNO NON È FARLI AVVICINARE MA NON FARLI ALLONTANARE PRECEDENTE

Figure A1. – An example of the individual learning journal. It is divided into four parts. The first one begins with the question: «What is today's learning goal?» and follows with «I must remember to pay attention to» (assessment criteria). The second part reports this sentence: «I tested my abilities and discovered that …»; then, it is possible to see three symbols representing the strengths (the star), the weaknesses (the balloon), and the improvement strategies (the light). The last part gives space to the exercises made using the improvement strategies.

APPENDIX B

Table B1. – Pre-test and Post-test B's characteristics. Data regarding the difficulty levels of the items was taken from the INVALSI technical reports, but these data refer to different populations and years (the translation was done by the authors of the contribution).

Pre-test and Post-test B's characteristics (items from INVALSI national standardized tests of previous years)			
Monitored reading comprehension abilities	*Narrative text*	*Expository text*	*Example of items*
Retrieve information that is explicitly stated in the text	1 (one correct answer)	1 (multiple choice question)	Where does the story you read take place?
Recognize concepts/ meanings expressed in paraphrastic forms (direct inferences or first-level inferences)	1 (multiple choice question)	1 (multiple choice question)	Considering what you read in the second paragraph, you can say that: A. all teens need to go to sleep at the same time; B. some tee ns need to sleep in a dark room; C. all teens need to sleep the same number of hours; D. some teens need to go to sleep before others.
Make lexical inferences	1 (edited) (multiple choice question)	1 (multiple choice question)	In the fourth paragraph the expression «remain in apnea» (line 54) means: A. remaining still; B. remaining on alert; C. remaining without breathing; D. staying asleep.
Make text-connecting inferences *Establish the reference of an anaphors or an anaphoric chain*	2 (multiple choice questions)	1 (one correct answer)	On line 13 there is the word «interlocutor». Who is the interlocutor you are talking about? A. The passerby. B. The young apprentice. C. Kannitverstan. D. The friend.
Make text-connecting inferences *Discover the implicit connections between the information of the text mainly using text's clues*	4 (three multiple choice questions and one with a single correct answer)	1 (multiple choice question)	«Good friend – he said – would tell me what the owner's name is of this beautiful house...?». When asking this question to the passer-by, the boy does not consider something that is important for what happens next. What is it? A. The passerby might not know the language he spoke. B. The passerby might not want to talk to a stranger. C. The passerby could have had an urgent task to attend to. D. The passerby might want to make a joke.
Total	9	5	

ECPS Journal – 31/2025 - https://www.ledonline.it/ECPS-Journal/
Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Table B2. – Post-test A's characteristics. The data regarding the difficulty levels of the items was taken from the INVALSI technical reports, but these data refer to different populations and years. Note that it's impossible to call this test «parallel» since the correlation between the measures obtained from Post-test B and Post-test A was not monitored before their administration to students (the translation was done by the authors of the contribution).

Post-test A's (parallel test) characteristics (items from INVALSI national standardized tests of previous years)			
Monitored reading comprehension abilities	*Narrative text*	*Expository text*	*Example of items*
Retrieve information that is explicitly stated in the text	1 (one correct answer)	1 (multiple choice question)	If you consider the information given in the first paragraph on the timing of the passive urbanization process, you can say that: A. it is a process that happens gradually, year after year; B. it is a very time-consuming process; C. it is a process that happens quickly, in a few days; D. it is a process that takes about two years.
Recognize concepts/ meanings expressed in paraphrastic forms (direct inferences or first-level inferences)	1 (multiple choice question)	1 multiple choice question)	«I must teach him to be a bear» (line 54). According to the text, this means that the boy wants to teach the bear: A. to know the forest and to get food for themselves; B. to be wary and aggressive towards the unfamiliar; C. to warm up by the fire without burning his legs; D. to hide in the thick of the forest not to be seen.
Make lexical inferences	1 (multiple choice question)	1 (multiple choice question)	The expression «It was the time of the games» (line 36) means that: A. bears play for a long time; B. it is the childhood of the bear; C. when it rains, the boy and the bear can play; D. to be friend of a bear you have to play with him.
Make text-connecting inferences *Establish the reference of an anaphors or an anaphoric chain*	2 (multiple choice questions)	1 (one correct answer)	«On the contrary, the most sensitive and specialized species meet various difficulties, and they may therefore decrease» (rows 35-36). This phrase leads to the question: «Unlike who or what?». Find the information you need to answer and type or write it in your own words. Unlike …
Make text-connecting inferences *Discover the implicit connections between the information of the text mainly using text's clues*	4 (multiple choice questions)	1 (multiple choice question)	«It was almost a year since Djidi had arrived in the village and that night, like all nights, Griska and the bear snuck out of the village». Why are the boy and the bear leaving the village? A. To show the tribe that they are brave. B. To experience the thrill of challenging adult rules. C. To test themselves and measure their cunning. D. To experience, learn and grow.
Total	9	5	

Riassunto

Valutazione formativa (in inglese «formative assessment» – FA) è un'espressione che copre diverse prassi. Se questa pluralità ha condotto allo sviluppo di diversi studi nel contesto internazionale, in Italia vi sono poche ricerche sperimentali sul tema. Pertanto, il contributo descrive una sperimentazione che esplora gli effetti dell'uso di pratiche di FA in classe sulle capacità di comprensione del testo degli studenti. Si tratta di un piano a due gruppi che ha coinvolto gli studenti di due classi prime di una scuola secondaria di primo grado, assegnati casualmente al gruppo sperimentale e a quello di controllo. Dopo aver svolto una prova di comprensione di testi, il ricercatore ha sviluppato 15 attività di FA con gli studenti del gruppo sperimentale. Al termine, sono state misurate nuovamente le abilità degli studenti ed è stata verificata la differenza pre- post-test tra i due gruppi. Durante l'analisi, sono stati eseguiti test non parametrici che non hanno rivelato differenze statisticamente significative. Nonostante ciò, una tendenza mostra un aumento leggermente più elevato del raggiungimento del gruppo sperimentale.

Parole chiave: Abilità di comprensione del testo; Disegno sperimentale; Effective teaching; Prassi di formative assessment; Scuola secondaria di primo grado.

How to cite this paper: Guasconi, E., & Vannini, I. (2025). Formative assessment practices for improving students' text comprehension abilities: An experiment in a lower secondary school in Italy [Prassi di *formative assessment* per promuovere le abilità di comprensione del testo: una sperimentazione nella scuola secondaria di primo grado in Italia]. *Journal of Educational, Cultural and Psychological Studies (ECPS)*, *31*, 153-182. https://doi.org/10.7358/ecps-2025-031-guva