Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 18, Issue 2, December 2024, pp. 351-374. ISSN: 1307-6086

Research Article

The Effects of Formative Assessment on Student Achievement and Attitudes in Math*

Çağla AKBAŞ[1], Ceyhun OZAN[2]

[1] Ministry of National Education, TOBB Efkan Ala İmam Hatip Middle School, Türkiye, caglaakbas27@gmail.com, http://orcid.org/0000-0002-1474-1681

[2] Atatürk University, Kâzım Karabekir Faculty of Education, Türkiye, ozanceyhun@atauni.edu.tr, http://orcid.org/0000-0002-1415-7258

Received : 08.07.2024 Accepted : 22.10.2024

Doi: https://doi.org/10.17522/balikesirnef.1512149

Abstract – The objective of formative assessment is to allocate or assign a value to the student, identify the student's learning deficiencies and make instructional arrangements to eliminate these deficiencies. This study employed a non-equivalent group pre-test and post-test design to investigate the effects of formative assessment on fifth-grade students' academic achievement and their attitudes toward mathematics. The independent variables were formative assessment and level-determining evaluation methods. The dependent variables were students' academic achievements and attitudes toward mathematics. The research group comprised fifth-grade students of a determined secondary school in the Pasinler district of Erzurum province. Through random assignment, 17 students were included in the experimental group and 13 in the control group. The data were obtained using the achievement test and attitude scale. This research was conducted in a secondary school fifth-grade math class for 11 weeks. The mathematics teacher, the researcher, delivered lessons to both groups. The mathematics achievement of students in the experimental group was higher than the control group. However, the difference was not statistically significant. Similarly, no significant difference was found in the attitudes of students toward math. Various recommendations were made in light of the findings obtained in this study.

Keywords: Formative assessment, mathematics education, mathematics achievement, attitude to mathematics, elementary education, secondary school.

Corresponding author: Ceyhun OZAN, ozanceyhun@atauni.edu.tr

* This study was produced from the master's thesis published by first author under the supervision of the second author. This paper was produced from the project supported by Atatürk University Scientific Research Projects Unit with the code SYL-2022-10227.

Introduction

Formative assessment is a method that entails collecting and utilizing data on student learning to offer continuous feedback and assistance toward achieving academic objectives. Black and Wiliam (1998a) characterize formative assessment as “the approach employed by educators and learners to identify and address student learning to enrich that learning throughout the instructional process.” Put differently, formative assessment strives to enhance the teaching and learning experience by conducting evaluations and providing feedback that aligns with learning objectives throughout the learning journey. If a student's understanding is not checked, it becomes impossible to know what they have learned from the lesson (Kültür, 2021).

Formative assessment is an essential tool for teachers and students. It enables instructors to recognize struggling students early, foster student engagement and motivation, and promote more profound learning. Therefore, it should be an essential component of any curriculum. By providing feedback on their performance and guiding them to identify areas where they require improvement, formative assessment can promote profound learning, boost student motivation and engagement, and facilitate realizing their maximum potential (Black & Wiliam, 1998a; Hattie & Timperley, 2007; Sadler, 1989).

While formative assessment is recognized as a valuable tool for learners at all stages, fifth grade represents a crucial period in students' academic development. Students’ transition from elementary foundational learning to more complex cognitive tasks at this stage, making it an ideal time to examine how formative assessment can support this critical phase. Research has shown that formative assessment is especially effective in promoting student engagement and improving outcomes during this developmental stage, where the demands of the curriculum increase and students are expected to demonstrate higher-order thinking skills (Black & Wiliam, 1998a). Therefore, focusing on fifth-grade students allows the study to address this critical juncture in learning, providing valuable insights into how formative assessment can facilitate academic achievement during a pivotal time in students' educational trajectories.

Features of Formative Assessment

As noted by Black et al. (2003), formative assessment comprises four fundamental elements: (i) Clarifying learning objectives and achievement criteria, (ii) Promoting the quality of questioning/dialogue, (iii) Improving the quality of scoring/feedback/record keeping (iv) Using peer and self-assessment. The "big idea" is to leverage insights into student learning to adapt instruction for enhanced responsiveness to individual student needs. In other words, instruction is adaptable to the student's learning needs.

Clarifying learning objectives and achievement criteria is fundamental to effective teaching and learning practices. Learning objectives refer to the goals students should achieve at the end of a lesson or unit, while criteria for achievement refer to the standards or expectations for achieving those objectives (Black & Wiliam, 1998a; Rodriguez & Albano, 2017). Hattie (2009) contends that the effective communication and clarification of learning objectives and achievement criteria for students represent a robust pedagogical strategy for improving student achievement. More recent studies, such as those by Brookhart (2017), have expanded on this by emphasizing that clear criteria improve academic outcomes and foster student autonomy.

Questioning and dialogue are essential components of formative assessment (Hodgson & Pyle, 2010; Walsh, 2022). Borich (2014) suggests that in a typical primary or secondary school class hour, 50 or more questions may be asked and that up to 80 percent of classroom time can be devoted to questioning and answering. Through meaningful dialogues, educators and learners can engage in an interactive exchange of ideas to resolve misunderstandings and promote a more profound level of learning (Black & Wiliam, 1998a; Heritage, 2007). Walsh and Sattes (2016) highlight that high-quality questioning promotes engagement and fosters metacognition, as students reflect on their thinking processes.

For formative assessment to be practical, improving the quality of scoring, feedback, and record-keeping is imperative. As Hattie and Timperley (2007) emphasize, feedback is a central formative assessment component. Andrade et al. (2015) contend that the efficacy of formative assessment depends on the quality of feedback given to students regarding their learning progress and teachers regarding their instructional practices. Feedback on formative assessment wields a significant influence on both learning and achievement. Some researchers even state that feedback is the most crucial element of students' learning (Bell, 2007; Brown, 2018; Shavelson et al., 2008; Wiliam, 2018).

Self and peer assessment are other effective formative assessment components (Nicol & Macfarlane-Dick, 2006). The responsibility for offering feedback is not exclusive to teachers. When properly instructed and guided, self and peer assessments can also provide students with valuable feedback focused on learning (Andrade, 2019; Andrade et al., 2015; Huisman et al., 2019; Panadero et al., 2018). Peer assessment or feedback involves students providing comments on the work of their peers (Topping, 2009).

Formative Assessment in Mathematics Education

The literature highlights that formative assessment is a powerful tool in mathematics education, as it facilitates active engagement and provides real-time feedback, both of which are essential for learning (Black & Wiliam, 1998b; Hattie & Timperley, 2007). However, a deeper exploration into what has been written specifically about formative assessment within the context of mathematics education is necessary. Research emphasizes that mathematics plays a pivotal role in developing critical problem-solving abilities and logical reasoning skills, which are crucial for academic and real-world success (Boaler, 2016; Schoenfeld, 2013). Given the abstract nature of many mathematical concepts, formative assessment provides a scaffold for students to make connections between different concepts and correct their misunderstandings before they become entrenched (Swan, 2006).

Despite its significance, many students may face significant challenges with mathematics, leading to poor academic performance, decreased self-efficacy, and lower motivation (Pajares & Graham, 1999; Wigfield & Eccles, 2000). Therefore, by providing regular feedback and helping teachers adjust instruction based on students' progress, formative assessment has been proven to support learning in this field (Black & Wiliam, 1998a; Wiliam, 2011).

In mathematics, formative assessment practices include a variety of strategies such as diagnostic questions, exit tickets, and real-time quizzes, which allow teachers to gather data about student progress. These tools help identify common misconceptions (Fennell, 2011) and provide an opportunity to differentiate instruction, ensuring that all students receive the support they need to succeed (Clarke, 2005). However, this study needs to further engage with what previous studies have specifically revealed about formative assessment in mathematics to underscore its contribution to the existing body of knowledge.

Studying formative assessment in mathematics is crucial because it addresses academic and affective challenges students face in this subject. The cyclical process of data collection, analysis, and instructional adjustment inherent in formative assessment (Black & Wiliam, 1998a) is particularly beneficial for mathematics, where frequent misconceptions can hinder students' understanding. Recognizing these misconceptions and addressing them promptly can significantly improve learning outcomes. Since mathematics is a cumulative discipline, early identification of gaps through formative assessment can prevent future difficulties and enhance students' problem-solving and logical reasoning abilities. The study's focus on the practical application of formative assessment techniques such as problem exploration, group assessments, and peer collaboration (Lee, 2006) shows how tailored feedback and instructional adjustments can lead to better outcomes in mathematics education. These techniques highlight formative assessment's potential to foster a deeper understanding of mathematics and its relevance to ongoing educational reforms.

While there is substantial research supporting the benefits of formative assessment for student achievement in mathematics (Black & Wiliam, 1998a; Hattie & Timperley, 2007; Heritage et al., 2009; Sadler & Good, 2006), this study adds value by addressing specific strategies and tools that can be used to identify students' learning gaps in real-time. Furthermore, by focusing on formative assessment's impact on students' self-efficacy (Schiefele et al., 2012), problem-solving skills (Cavanagh et al., 2016; Kramarski & Michalsky, 2009; Beatty & Gerace, 2009), and motivation (Hounsell et al., 2008), this study extends understanding of how formative assessment can influence not only academic outcomes but also students' emotional and cognitive engagement with mathematics. The present study emphasizes that formative assessment is not just about improving test scores but about fostering a growth mindset and creating a more inclusive learning environment that supports diverse learners. The findings of this research will help bridge the gap between theory and practice, offering educators practical tools to enhance students' learning experiences in mathematics classrooms.

Research Questions

This research investigates the influence of fifth-grade students' formative assessment on mathematics achievements and attitudes toward mathematics in secondary schools. To address this objective, this study outlines the following research questions:

1) Is there a significant difference between the control group's pre-test and post-test mathematics achievements?
2) Is there a significant difference between the experimental group's pre-test and post-test mathematics achievements?
3) Is there a significant difference between the control group's pre-test and post-test attitudes toward math?
4) Is there a significant difference between the experimental group's pre-test and post-test attitudes toward math?

5) Is there a significant difference between the experimental and control groups' post-tests?

Method

Research Design

In this research, we utilized a non-equivalent group pre-test and post-test design. This design is classified as experimental because it allows for comparing outcomes between an experimental group that receives the intervention (formative assessment) and a control group that does not. According to Campbell et al. (1963), pre-test and post-test measures help establish a baseline for assessing the effectiveness of the intervention while controlling for pre-existing differences between groups.

Due to logistical constraints, it was not feasible to randomly assign individuals to either the experimental or control groups. However, we made every effort to ensure that participants were assigned to the groups as randomly as possible, thereby minimizing potential biases and confounding factors (Shadish et al., 2002). This approach aligns with the principles outlined by Creswell (2014), emphasizing the importance of quasi-experimental designs when random assignment is impractical, yet the need for causal inferences remains.

Study Group

The study group consisted of fifth-grade students enrolled in a secondary school located in the Pasinler district of Erzurum province, Türkiye. The selection of the study group was based on the researcher's accessibility and the possibility of attending mathematics classes and extracurricular activities with the group on an individual basis. The researcher's presence throughout the application process enabled the observation of students' interests, attitudes, and behaviors. Although the study group's district is not particularly large, it can be described as having a low-to-middle income level in general. Most students in the study group have a middle-income level, while their parents' education levels are mainly at the primary school level. The average class size throughout the county is currently 16. A total of 30 students participated in this study, with 17 from the 5/B class and 13 from the 5/C class in the school representing fifth-grade students in the county. Demographics of the study group are shown in Table 1.

Table 1 Demographics of Study Group

Gender	Control	Experimental	Total
Female	5	10	15
Male	8	7	15
Total	13	17	30

The results regarding whether there was a significant difference between the mathematics achievement and attitude scores of the experimental and control groups before the experimental procedures are shown in Table 2.

Table 2 Results of Pre-tests

Group	n	$\bar{X}$	SS	sd	t	p
Control (achievement)	13	4.61	1.98	28	1.830	.078
Experimental (achievement)	17	5.88	1.79			
Control (attitude)	13	3.78	.37	28	2.798	.009
Experimental (attitude)	17	4.22	.49			

As shown in Table 2, it was determined that there was no significant difference between the mathematics achievement scores of the experimental and control group students ($t(28) = 1.830$; $p > .05$). A significant difference was found in the mathematics attitude scores between the experimental and control groups ($t = 2.798$; $p < .05$). The attitude scores of the experimental group were significantly higher before the experimental procedures.

Process

This research was conducted over the course of 11 weeks in a fifth-grade mathematics class at a secondary school. The lessons were taught by the mathematics teacher, the researcher, to both groups. It should be noted that formative assessment practices were only applied in the experimental group, while no such practices were implemented in the control group. The lessons in the experimental group were designed with the intention of achieving the same learning outcomes, while also incorporating formative assessment applications. The teacher who conducted the practices in this study is a graduate student. Before the experimental procedures began, the teacher received weekly guidance from the thesis advisor on formative assessment practices. This training lasted one hour per week throughout the entire semester, ensuring that the teacher had a clear understanding of how to implement formative assessment techniques during this study effectively. However, it is recommended that more comprehensive and structured professional development programs be provided to teachers before such studies to enhance their capacity to apply formative assessment effectively.

Black et al. (2003) identified four key elements of formative assessment: (i) the clarification of learning objectives and criteria for achievement, (ii) the promotion of the quality of questioning and dialogue, (iii) the improvement of the quality of scoring and feedback, and (iv) the utilisation of peer and self-assessment. The teaching process of Ozan and Kıncal's (2018) research included the following practices within the framework of the four fundamental components of formative assessment:

The Clarification of Learning Objectives and Criteria for Achievement

This study employed the following strategies to elucidate the learning objectives and achievement criteria: (i) The teacher introduced the learning objectives of each lesson and engaged in a discussion with the students about the expected learning outcomes. (ii) The teacher frequently reminded the students of the learning objectives throughout the lesson. (iii) After completing the lesson, the teacher discussed with the students what they had learned and how it related to the learning objectives. (iv) The criteria for achievement in the activities that students would undertake in the classroom were made clear, including the necessary actions and the means of attaining them. (v) The criteria for completing homework assignments successfully were also made known to the students.

The Promotion of the Quality of Questioning and Dialogue

This study employed the following strategies to enhance the quality of questioning and dialogue: (i) Collaborative group work was employed to facilitate student dialogue. (ii) The teacher formed groups heterogeneously, considering factors, such as gender, academic performance, and affective traits of the students. The study groups consisted of four groups with four students each and one group with five students. (iii) At the conclusion of each unit, the students were assigned to new groups through random selection. Students were encouraged to provide feedback to each other during group work. (iv) The teacher employed high-level thinking skill questions throughout the course. (v) Students were allotted time to reflect before answering questions, with the amount of reflection time varying from 3 to 25 seconds, based on the complexity of the question. (vi) As part of assessing their higher-level thinking abilities, students were given the opportunity to engage in peer discussions. (vii) The "No Hands Up" strategy was employed to the teacher's questions. This strategy aimed to enhance the participation and engagement of all students by prohibiting them from raising their hands, thereby requiring each student to contemplate the answer and participate in the lesson.

The Improvement of the Quality of Scoring and Feedback

This study employed the following strategies to enhance the quality of scoring, feedback: (i) Instead of assigning grades or points, the teacher primarily provided students with immediate feedback on their classwork or homework. (ii) Students were given the opportunity to amend their work in accordance with the feedback provided by the teacher. (iii) Quizzes were administered at the conclusion of each topic and unit to identify students' learning gaps and provide feedback. (iv) Students were provided with feedback on their weaknesses following the quiz. (v) Based on the feedback students received on their quizzes, assignments, and activities, adjustments were made in the classroom. Many arrangements were made to address the identified weaknesses, including small group work, individual activities, re-explaining, worksheets, teaching by showing and doing, internet research, clarification of assessment criteria, additional reading, and concept mapping. Efforts were made to create a learning-oriented classroom culture with student-centered practices in which students are active. (vi) Both summative and formative assessment data were recorded when assessing students. Student names were converted into codes. (vii) The assessment data on the students' performance were shared with the following year's teacher to ensure continuity in their learning. (viii) Rewards were given to recognize and encourage students' progress. (ix) The awards were not used to reward students by comparing them with each other but by considering their development levels. Rewards were provided to students who demonstrated individual progress in quizzes.

The Utilization of Peer and Self-assessment

This study incorporated the following methods for utilizing self and peer assessment: (i) The regular integration of self and peer assessments into the lessons. Self-assessment and peer assessment activities were carried out after each unit. (ii) Guidelines were presented for each activity, and students were educated on performing self and peer assessments. (iii) The completed assignments were discussed among the students to assess whether they met the expected standards.

The control group’s instructional sessions followed the standard curriculum without incorporating formative assessment strategies. As a result, several key elements typically associated with formative assessment were absent. For example, no cooperative group work was facilitated, and high-level thinking questions were rarely posed. Additionally, there were no short quizzes to gauge student understanding regularly, no awards or incentives were provided to motivate student performance, and self-assessment and peer assessment activities were not part of the instructional process. This traditional approach focused on delivering the material as per the curriculum guidelines without the additional scaffolding formative assessment provides.

Data Collection Tools

Mathematics Achievement Test

The Mathematics Achievement Test was used to measure the mathematical achievement of students in the experimental and control groups. The test was developed based on the learning outcomes delineated in the "Mathematics Curriculum," published by the Republic of Türkiye Ministry of National Education (MoNE, 2017). The test encompassed 37 questions about 17 learning outcomes. The opinions of three mathematics educators from the same educational institution and a curriculum and instruction expert were sought to ensure the test's reliability. Following careful consideration, the experts concluded that no changes were required. The test was administered to sixth-grade students who had received mathematics education the previous year. Item analysis was performed after a preliminary application of the test to 22 sixth-grade students, and item difficulty and discrimination levels were determined. We created 27 percent lower and upper groups to calculate item discrimination. Table 3 shows the item difficulty and discrimination indexes of the pre-application test.

Table 3 Pre-application of the Achievement Test's Degree of Item Difficulty and Discrimination

Item number	Difficulty index	Discrimination index	p	Item number	Difficulty index	Discrimination index	p
1	0.90	-0.16	.25	20	0.13	-0.16	.35
2*	0.63	0.66	.00	21*	0.68	0.66	.00
3	0.50	0.33	.00	22	0.54	0.33	.00
4*	0.77	0.5	.00	23	0.31	0.16	.28
5	0.22	0.33	.00	24*	0.50	0.33	.00
6*	0.31	0.33	.00	25	0.45	0.33	.00
7*	0.27	0.33	.00	26*	0.09	0.33	.00
8	0.04	0.16	.30	27	0.18	-0.16	.30
9*	0.18	0.50	.00	28	0.27	0	.40
10	0.50	0.33	.00	29*	0.59	0.50	.00
11*	0.63	0.33	.00	30	0.40	0.16	.20
12	0.22	0.16	.27	31	0.54	0.16	.15
13*	0.27	0.66	.00	32	0.63	0	.50
14	0.22	0	.40	33*	0.50	0.50	.00
15*	0.68	0.83	.00	34*	0.72	0.50	.00
16	0.54	0.50	.00	35	0.77	0.50	.00
17	0.77	0.33	.00	36	0.81	0.50	.00
18*	0.22	0.50	.00	37*	0.72	0.66	.00
19*	0.5	0.50	.00				

*Items taken to the final test

According to Table 3, the degree of item discrimination varied from 0.16 to 0.83. Items with discrimination of less than 0.30 were excluded from the test. When selecting one question from each learning outcome based on their high degrees of discrimination, the discrimination values of the 17 selected items ranged from 0.33 to 0.83. The mean discrimination of the 17 selected items was 0.50. Each item in the test was assigned a point value, and the total score ranged from 0 to 17. The reliability of the achievement test was assessed by calculating the KR-20 reliability coefficient, yielding a value of 0.71. Tests with a KR-20 coefficient greater than 0.7 were considered reliable. Therefore, the achievement test is reliable.

Attitude Scale toward Mathematics

The assessment of shifts in students' attitudes toward mathematics before and after the intervention relied on the utilization of "The Attitude Scale towards Mathematics" (ASTM), formulated by Önal (2013). This scale comprises 22 items and four factors: anxiety, interest, necessity, and work. There are 11 negative items on ASTM. The scale items are presented utilizing a 5-point Likert scale format, where respondents can choose from options such as "strongly disagree," "disagree," "undecided," "agree" and "strongly agree."

The overall scale demonstrated a Cronbach's alpha of 0.90. The Cronbach's alpha coefficients of the dimensions of anxiety (5 items), interest (10 items), necessity (3 items), and working (4 items) were 0.74, 0.89, 0.70, and 0.69, respectively (Önal, 2013). In this research, Cronbach's alpha for the overall scale was 0.82.

Data Analysis

The normality of the data collected from the control and experimental groups was assessed using the Shapiro-Wilk normality test. The Shapiro-Wilk test is known for its high sensitivity in detecting deviations from normality, making it particularly effective for small sample sizes (Razali & Wah, 2011). Table 4 shows the Shapiro-Wilk results.

Table 4 Shapiro-Wilk Results

Test	Group	Statistic	p
Mathematics achievement pre-test	Control	0.960	.75
	Experimental	0.939	.31
Mathematics achievement post-test	Control	0.894	.11
	Experimental	0.954	.52
Mathematics attitude pre-test	Control	0.949	.58
	Experimental	0.910	.10
Mathematics attitude post-test	Control	0.922	.26
	Experimental	0.960	.63

The Shapiro-Wilk test results indicated that both the control and experimental groups exhibited normality in their mathematics achievement and attitude scores for both pre-tests and post-tests, as all p-values were more significant than 0.05. After confirming that the collected data showed normal distribution and had equal variance, we used independent samples t-test to compare the mean differences between the control and experimental groups in terms of mathematics achievement and attitudes toward mathematics and analysis of covariance (ANCOVA) to compare the mean differences between the control and experimental groups post-tests. ANCOVA helps assess the effect of the treatment while accounting for the initial differences in pre-test scores, thereby providing a clearer understanding of the impact of formative assessment on the students' achievements and attitudes (Tabachnick & Fidell, 2019).

Threats to Internal Validity in Experimental Designs

In this study, the authors acknowledge that the research design may be considered weak due to the use of non-equivalent groups. Non-equivalent groups can introduce threats to internal validity, including selection bias, maturation effects, and history effects. To address these potential threats and enhance the robustness of the findings, the authors implemented several strategies:

1. Selection Bias: Although random assignment to groups was not feasible due to logistical constraints, efforts were made to ensure that participants were assigned to groups as randomly as possible within the given classes. This minimizes the likelihood that pre-existing differences between the groups would confound the results.

2. Maturation Effects: To control potential maturation effects, the study was conducted over a relatively short duration (11 weeks). Additionally, pre-test scores were used to assess the initial equivalence of groups, allowing for a more accurate comparison of post-test outcomes.

3. History Effects: The authors took measures to reduce the impact of external events that might influence student learning. By conducting the study in a controlled classroom environment and maintaining consistent teaching practices across both groups, the authors aimed to limit the influence of outside factors on the students' achievement and attitudes toward mathematics.

4. Regular Monitoring and Feedback: Throughout this study, regular assessments were conducted to monitor student progress and provide feedback. This ongoing evaluation helped identify any unforeseen issues affecting internal validity, allowing for timely interventions.

Findings

Table 5 shows descriptive statistics regarding the scores of both groups, along with the pre-test and post-test scores derived from the Achievement Test and ASTM.

Table 5 Descriptive Statistics

	Groups	n	Mean	Adjusted mean
Achievement (Pre-test)	Control	13	4.61	
	Experimental	17	5.88	
ASTM (Pre-test)	Control	13	3.78	
	Experimental	17	4.22	
Achievement (Post-test)	Control	13	7.61	7.69
	Experimental	17	9.58	9.01
ASTM (Post-test)	Control	13	3.75	3.99
	Experimental	17	4.09	4.00

As indicated in Table 5, the control group students exhibited a pre-test achievement score of 4.61, which showed an increase to 7.61 in the post-test. In contrast, the experimental group students started with a pre-test achievement score of 5.88, demonstrating an improvement to 9.58 in the post-test. In terms of the ASTM scores, the control group students began with a pre-test score of 3.78, rising to 7.61 in the post-test.

There was a significant difference in the mathematics achievement scores among the control group students in favor of the post-test ($t(12) = -2.98$; $p < .05$). There was no significant difference in the attitude scores toward mathematics among the control group students ($t(12) = 2.67$; $p > .05$).

There was a significant difference in the mathematics achievement scores among the experimental group students in favor of the post-test ($t(16) = -5.370$; $p < .05$). There was no significant difference in the attitude scores toward mathematics among the experimental group students ($t(16) = 1.284$; $p > .05$). The statistical significance of the difference in students' post-test achievement scores was investigated using ANCOVA, with the results detailed in Table 6.

Table 6 Post-test Scores of Achievements

Source	Sum of squares	df	Mean square	F	p
Corrected Model	85.920	3	28.640	2.979	.050
Intercept	88.559	1	88.559	9.212	.005
Pre-test	32.858	1	32.858	3.418	.076
Group	10.810	1	10.810	1.125	.299
Error	249.946	26	9.613		
Total	2624.000	30			
Corrected total	335.867	29			

Table 6 shows no significant difference between the post-test achievement scores of the groups ($F_{(1, 26)} = 1.125$; $p > 0.05$). As a result, the implemented procedures did not yield a statistically significant difference in the achievement of the experimental group students compared to the control group students. Table 7 shows the ANCOVA results for ASTM post-test scores.

Table 7 Post-test Scores of Attitudes

Source	Sum of squares	df	Mean square	F	p
Corrected model	4.176	3	1.392	8.773	000
Intercept	.305	1	.305	1.925	.177
Pre-test	2.757	1	2.757	17.377	.000
Group	.271	1	.271	1.705	.203
Error	4.125	26	.159		
Total	476.017	30			
Corrected total	8.301	29			

There was no significant difference in the post-test scores regarding the attitude toward mathematics between the groups ($F_{(1, 26)} = 1.705$, $p > 0.05$).

Conclusion, Discussion and Recommendations

In the present study, the observed difference was not statistically significant despite the experimental group, where formative assessment interventions were implemented, demonstrating higher mathematics achievement than the control group. In the literature, several studies have shown that formative assessment and feedback positively affect students' achievement. These studies include systematic reviews or meta-analyses studies, such as those by Black and Wiliam (1998a), Hattie and Timperley (2007), Shute (2008), and Kluger and DeNisi (1996). Moreover, recent studies have also focused on the effects of formative assessment on specific skills and subjects (Bulunuz et al., 2017; Gedikli, 2018; Kıncal & Ozan, 2018; Kültür, 2021; Ozan & Kıncal, 2018; Sönmez, 2020; Unaş, 2021).

In this study, the authors recognize that the absence of a significant difference in student achievement and attitudes toward mathematics may be attributed to several factors. First, the literature reveals that some studies indicate no significant effect of formative assessment on student achievement (Andrews, 2011; Collins, 2012; King, 2003; Tuominen, 2008; Yin et al., 2008). This suggests that the effectiveness of formative assessment can vary depending on various contextual and instructional factors.

One major challenge in the effective implementation of formative assessment is its integration into classroom instruction. Successful enactment requires that teachers understand effective formative assessment practices and have access to the necessary resources and support. The authors emphasize that teachers must receive adequate training to utilize formative assessment effectively (Black & Wiliam, 1998b; Earl, 2003; Heritage, 2007; Popham, 2008; Ruiz-Primo et al., 2008).

Furthermore, implementing formative assessment can demand additional time and effort from teachers, which may not always be feasible within the existing curriculum constraints. For example, the authors noted that for students with lower overall achievement to benefit from formative assessment fully, the time allocated to subjects might need to be increased.

In addition, the lack of access to essential technologies, such as smart boards and internet resources in classrooms, may further hinder the development of effective formative assessment practices. Without these tools, teachers may struggle to create engaging and interactive learning environments conducive to formative assessment, which could explain the lack of significant differences observed in the study's findings. By highlighting these challenges, the authors underscore the complexity of implementing formative assessment in educational settings and suggest that further research is needed to explore these dynamics.

Despite incorporating formative assessment practices within the experimental group in this study, no statistically significant difference was observed in the mathematics achievement of students when compared to the control group. Similarly, no significant difference was noted in the attitudes of students in the experimental group toward math than those in the control group. It is crucial to recognize that the pivotal components of formative assessment encompass feedback and self-assessment. In line with the feedback given to students within the scope of formative assessment, students must make various arrangements and try to guide themselves in line with their self-assessment. Students need time to develop these skills. Achievement did not increase at a statistically significant level in this study because the students needed to follow the regulations adequately after the feedback given by the teacher or the duration of one semester required to be increased. Koçak (2021) conducted a study with geography teachers and found that while teachers' attitudes toward formative assessment were positive, there were deficiencies in theoretical and practical terms, and students could not be sufficiently motivated to this assessment without grade concern. The existing education system's focus on result-oriented assessment approaches was also stated as a difficulty in practice.

The math curriculum issued by the Ministry of National Education in Türkiye (MoNE, 2017) includes only cognitive gains, and it is challenging to achieve affective gains for math in a curriculum that does not give place to affective gains. Students in Türkiye mostly think that mathematics is a complex subject, which creates anxiety and a negative attitude toward the subject. Baykul (2005) suggests that changing students' attitudes towards math is not straightforward and is a long-term process. Therefore, implementing formative assessment practices over an extended period and not only for certain subjects but also for all the subjects in the curriculum with a formative approach may be more effective in changing students' attitudes toward the lesson.

The main shortcomings of formative assessment are time and resource limitations (Black & Wiliam, 1998a; Brookhart, 2013; Heritage, 2010; McMillan & Hearn, 2008; Popham, 2011; Sadler, 1998). Formative assessment can be dysfunctional when teachers cannot devote sufficient time to student feedback. In this case, teachers may be unable to adequately assess students' performance or provide feedback. Another notable limitation is the teachers' insufficient knowledge and experience regarding formative assessment methods. In this case, teachers may not assess students correctly or give incorrect feedback (Chappuis, 2009).

This study highlights the critical role of formative assessment in enhancing students' learning outcomes and improving teaching practices. Formative assessment has the potential to elevate student performance by providing real-time feedback and insights into their understanding. However, this study found no statistically significant differences between groups, suggesting that the effectiveness of formative assessment may depend on various contextual factors.

The absence of statistically significant differences in this study may stem from several factors. First, the limited training and teachers' knowledge in implementing effective formative assessment strategies can hinder its impact. Additionally, logistical constraints, such as insufficient time for providing feedback and inadequate resources like access to technology, may have affected the implementation process. Moreover, the small sample size and the non-equivalent nature of the groups could also limit the generalizability of the findings.

Several limitations may influence the study's results. The non-equivalent group design poses a threat to internal validity, as it does not control for pre-existing differences between the experimental and control groups. The reliance on self-reported data for attitudes toward math may also introduce bias. Furthermore, the researcher’s dual role as the teacher may lead to potential biases in assessment and feedback. Finally, external factors, such as classroom dynamics and students' home environments, were not controlled for, which could have impacted the outcomes.

For further research, it is recommended that studies explore the long-term effects of formative assessment on student achievement and attitudes across diverse educational contexts. Additionally, investigating the impact of teacher training programs on the effective implementation of formative assessment could provide valuable insights. Further studies should also consider using a larger sample size and random assignment to strengthen the validity of the findings. Finally, examining the role of technology in facilitating formative assessment practices could be beneficial, particularly in addressing the challenges of accessibility and engagement.

In conclusion, while formative assessment is a powerful tool for enhancing learning and teaching, its effectiveness is contingent upon proper implementation, adequate training, and supportive resources. By addressing the identified limitations and building on the study's findings, educators and researchers can better harness the potential of formative assessment in educational settings.

Compliance with Ethical Standards

Disclosure of potential conflicts of interest

No conflict of interest.

Funding

None.

CRediT author statement

The article was collaboratively written by two authors, with each contributing equally to its content.

Research involving Human Participants and/or Animals

The ethics committee permissions were obtained with the letter dated 11.11.2021 and numbered 12/10 of the "Social and Human Sciences Scientific Research and Publication Ethics Committee of University Institute of Educational Sciences."

Matematik Eğitiminde Biçimlendirici Değerlendirmenin Öğrencilerin Akademik Başarı ve Tutumlarına Etkisi

Özet:

Biçimlendirici değerlendirmenin amacı öğrenciye bir değer biçmek değil, öğrencinin öğrenme eksikliklerini belirlemek ve bu eksiklikleri gidermeye yönelik öğretimsel düzenlemeler yapmaktır. Araştırmanın temel amacı, matematik öğretimi sürecinde biçimlendirici değerlendirmenin beşinci sınıf öğrencilerinin akademik başarıları ve matematik dersine yönelik tutumlarına etkisini incelemektir. Bağımsız değişken, biçimlendirici değerlendirme ile düzey belirleyici değerlendirme uygulamalarıdır. Bağımlı değişkenler ise öğrencilerin akademik başarıları ve matematiğe karşı tutumlarıdır. Araştırmanın çalışma grubunu Erzurum ili Pasinler ilçesinde belirlenen bir ortaokulun beşinci sınıf öğrencileri oluşturmaktadır. Yansız atama yoluyla deney grubunda 17, kontrol grubunda 13 öğrenci yer almıştır. Verilerin elde edilmesinde başarı testi ve tutum ölçeği kullanılmıştır. Çalışma, ortaokul beşinci sınıf matematik dersinde 11 hafta boyunca yürütülmüştür. Dersler hem kontrol hem de deney gruplarında aynı zamanda araştırmacı olan matematik öğretmeni tarafından verilmiştir. Araştırma sonuçlarına göre, biçimlendirici değerlendirme uygulamalarının olduğu deney grubundaki öğrencilerin matematik dersi başarıları, kontrol grubundaki öğrencilere göre yüksek bulunmuştur ancak elde edilen fark istatistiksel olarak anlamlı değildir. Benzer şekilde, kontrol ve deney grubunun matematik dersi tutumlarında da anlamlı bir fark bulunmamıştır. Araştırma sonucunda çeşitli önerilerde bulunulmuştur.

Anahtar kelimeler: Biçimlendirici değerlendirme, matematik eğitimi, matematik başarısı, matematik tutumu, ortaokul eğitimi, ortaokul.

References

Andrade, H., & Lui, A., Palma, M., & Hefferen, J. (2015). Formative assessment in dance education. *Journal of Dance Education, 15(2)*, 47-59. https://doi.org/10.1080/15290824.2015.1004408

Andrade, H. L. (2019). A critical review of research on student self-assessment. *Frontiers in Education*, *4*, 1-13. https://doi.org/10.3389/feduc.2019.00087

Andrews, T. L. (2011). *The use of goal setting and progress self-monitoring with formative assessment in community college to increase academic achievement and self-efficacy* [Unpublished doctoral dissertation]. Temple University.

Baykul, Y. (2005). *Teaching mathematics in primary education (for grades 1–5)*. Pegem Academy.

Beatty, I. D., & Gerace, W. J. (2009). Technology-enhanced formative assessment: A research-based pedagogy for teaching science with classroom response technology. *Journal of Science Education & Technology*, *18*(2), 146. http://doi.org/10.1007%2Fs10956-0089140-4

Bell, B. (2007). Classroom assessment of science learning. In S. Abell, & N. Lederman (Eds.), *Handbook of research on science education* (pp. 965–1006). Lawrence Erlbaum.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Open University.

Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education*, *5*, 7-74. https://doi.org/10.1080/0969595980050102

Black, P., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *80*(2), 139-144. https://kappanonline.org/inside-the-black-box-raising-standards-through-classroom-assessment/

Boaler, J. (2016). *Mathematical mindsets: Unleashing students' potential through creative math, inspiring messages, and innovative teaching*. Jossey-Bass.

Borich, G. A. (2014). *Effective teaching methods: Research-based practice* (8th ed.).

Brookhart, S. M. (2013). *How to assess higher-order thinking skills in your classroom*. ASCD.

Brookhart, S. M. (2017). *How to give effective feedback to your students*. ASCD.

Brown, G. T. L. (2018). *Assessment of student achievement*. Routledge.

Bulunuz, M., Kıryak, Z., Tomaç, B., Karagöz, F., & Recepoğlu, B. (2017). Evaluation of formative assessment-based teaching practices: An action research. *Turkish Journal of Teacher Education 6(2),* 115-123. http://tujted.com/makale/630

Campbell, D. T., Stanley, J. C., & Gage, N. L. (1963). *Experimental and quasi-experimental designs for research.* Houghton, Mifflin and Company.

Cavanagh, A. J., Aragón, O. R., Chen, X., Couch, B., Durham, M., Bobrownicki, A., Hanauer, D. I., & Graham, M. J. (2016). Student buy-in to active learning in a college science course. *CBE Life Sciences Education, 15*(4), ar76. https://doi.org/10.1187/cbe.16-07-0212

Chappuis, J. (2009). *Seven strategies of assessment for learning.* Pearson.

Clarke, S. (2005). *Formative assessment in the secondary classroom*. Hodder Education.

Collins, N. M. (2012). *The impact of assessment for learning: Benefits and barriers to student achievement* [Unpublished doctoral dissertation]. Cardinal Stritch University.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage.

Earl, L. M. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Corwin.

Fennell, F. (2011). *Achieving fluency: Special education and mathematics*. National Council of Teachers of Mathematics.

Gedikli, H. (2018). *The effects of formative assessment practices in science education on students' metacognitive knowledge and skills* (Publication No. 503450) [Master's thesis, Sivas Cumhuriyet University]. Council of Higher Education Thesis Center.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses related to achievement.* Routledge.

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

Heritage, M. (2007). Formative assessment: What do teachers need to know and do?. *Phi Delta Kappan*, *89*(2), 140-145. https://kappanonline.org/formative-assessment-heritage/

Heritage, M. (2010). *Formative assessment: Making it happen in the classroom*. Corwin.

Heritage, M., Kim, J., Vendlinski, T. P., & Herman, J. L. (2009). From evidence to action: A seamless process in formative assessment?. *Educational Measurement, 28*(3), 24–31. https://doi.org/10.1111/j.1745-3992.2009.00151.x

Hodgson, C., & Pyle, K. (2010). *A literature review of assessment for learning in science*. National Foundation for Educational Research.

Hounsell, D., McCune, V., Hounsell, J., & Litjens, J. (2008). The quality of guidance and feedback to students. *Higher Education Research and Development, 27*(1), 55-67. https://doi.org/10.1080/07294360701658765

Huisman, B., Saab, N. van den Broek, P., & van Driel, J. (2019). The impact of formative peer feedback on higher education students' academic writing: A meta-analysis. *Assessment & Evaluation in Higher Education, 44*(6), 863-880, https://doi.org/10.1080/02602938.2018.1545896

Kıncal, R. Y., & Ozan, C. (2018). Effects of formative assessment on prospective teachers' achievement, attitude and self-regulation skills. *International Journal of Progressive Education, 14*(2), 77-92. https://doi.org/10.29329/ijpe.2018.139.6

King, M. D. (2003). *The effects of formative assessment on student self-regulation, motivational beliefs, and achievement in elementary science* [Unpublished doctoral dissertation]. George Mason University.

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. https://doi.org/10.1037/0033-2909.119.2.254

Koçak, B. (2021). *Determining the attitudes, perceptions and practices of geography teachers towards formative assessment* (Publication No. 676619) [Master's thesis, Necmettin Erbakan University]. Council of Higher Education Thesis Center.

Kramarski, B., & Michalsky, T. (2009). Investigating preservice teachers' professional growth in self-regulated learning environments. *Journal of Educational Psychology, 101*(1), 161–175. https://doi.org/10.1037/a0013101

Kültür, Y. Z. (2021). *The effect of formative assessment on secondary school students' mathematics course achievement and attitudes* (Publication No. 668735) [Master's thesis, Çukurova University]. Council of Higher Education Thesis Center.

Lee, C. (2006). *Language for learning mathematics.* Open University.

McMillan, J. H., & Hearn, J. (2008). Student self-assessment: The key to stronger student motivation and higher achievement. *Educational Horizons, 87*(1), 40-49. https://www.jstor.org/stable/42923742

MoNE, (2017). *Ortaokul matematik dersi (1.,2., 3., 4., 5, 6, 7 ve 8. sınıflar) öğretim programı*.

Nicol, D., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*, 199-218. https://doi.org/10.1080/03075070600572090

Ozan, C., & Kıncal. R. Y. (2018). The effects of formative assessment on academic achievement, attitudes toward the lesson, and self-regulation skills. *Educational Sciences-Theory & Practice, 18*, 85-118. https://doi.org/10.12738/estp.2018.1.0216

Önal, N. (2013). A study on the development of a middle school students' attitudes towards mathematics scale. *Elementary Education Online, 12*(4), 938-948. https://ilkogretim-online.org/index.php/pub/article/view/6042

Pajares, F., & Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemporary Educational Psychology, 24*(2), 124-139. https://doi.org/10.1006/ceps.1998.0991

Panadero, E. (2018). Is it safe? Social, interpersonal, and human effects of peer assessment: A review and future directions. In G. T. L. Brown, & L. R. Harris (Eds.), *Handbook of human and social conditions in assessment* (pp. 247-266). Routledge.

Popham, W. J. (2008). *Transformative assessment*. ASCD.

Popham, W. J. (2011). *Classroom assessment: What teachers need to know* (6th ed.). Allyn & Bacon.

Razali, N., & Wah, Y. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics, 2*(1), 21-33. http://www.de.ufpb.br/~ulisses/disciplinas/normality_tests_comparison.pdf

Rodriguez, M. C., & Albano, A. D. (2017). *The college instructor's guide to writing test items: Measuring student learning*. Routledge.

Ruiz-Primo, M. A., Briggs, D., Iverson, H., Talbot, R., & Shepard, L. A. (2011). Impact of undergraduate science course innovations on learning. *Science*, *331*, 1269-1270. https://doi.org/10.1126/science.1198976

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119-144. https:/doi.org/10.1007/BF00117714

Sadler, D. R. (1998). Formative assessment: revisiting the territory. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 77-84. https://doi.org/10.1080/0969595980050104

Sadler, P. M., & Good, E. (2006). The impact of self- and peer-grading on student learning. *Educational Assessment*, *11*(1), 1-31. https://doi.org/10.1207/s15326977ea1101_1

Schiefele, U., Schaffner, E., Möller, J., & Wigfield, A. (2012). Dimensions of reading motivation and their relation to reading behavior and competence. *Reading Research Quarterly*, *47*(4), 427-463. https://doi.org/10.1002/rrq.030

Schoenfeld, A. H. (2013). *Mathematical problem solving*. Elsevier.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Houghton, Mifflin and Company.

Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E. M., Ruiz-Primo, M. A., Tomita, M. K., & Yin, Y. (2008). On the impact of curriculum-embedded formative assessment on learning: A collaboration between curriculum and assessment developers. *Applied Measurement in Education*, *21*(4), 295-314. https://doi.org/10.1080/08957340802347647

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, *78*(1), 153-189. https://doi.org/10.3102/0034654307313

Sönmez, M. (2020). *The effect of formative assessment on reading comprehension* (Publication No. 649617) [Master's thesis, Düzce University]. Council of Higher Education Thesis Center.

Swan, M. (2006). *Collaborative learning in mathematics: A challenge to our beliefs and practices*. National Institute of Adult Continuing Education.

Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.

Topping, K. J. (2009). Peer assessment. *Theory into Practice, 48*(1), 20-27. https://doi.org/10.1080/00405840802577569

Tuominen, K. R. (2008). *Formative assessment and collaborative teaming with support involving middle school mathematics teachers* [Unpublished doctoral dissertation]. California University.

Unaş, F. (2021). *The effects of formative assessment practices in science education on students' self-regulated learning skills* (Publication No. 693052) [Master's thesis, Sivas Cumhuriyet University]. Council of Higher Education Thesis Center.

Walsh, J. A. (2022). *Questioning for formative feedback: Meaningful dialogue to improve learning*. ASCD.

Walsh, J. A., & Sattes, B. D. (2016). *Questioning for classroom discussion: Purposeful speaking, engaged listening, deep thinking*. ASCD.

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68-81. https://doi.org/10.1006/ceps.1999.1015

Wiliam, D. (2011). What is assessment for learning?. *Studies in Educational Evaluation, 37*(1), 3-14. https://doi.org/10.1016/j.stueduc.2011.03.001

Wiliam, D. (2018). Feedback: At the hard of -but definitely not all of- formative assessment. In A. A. Lipnevich, & J. K. Smith (Eds.), *The Cambridge handbook of instructional feedback* (pp. 3-28). Cambridge University.

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. https://doi.org/10.1080/08957340802347845