International Journal of Instruction
e-ISSN: 1308-1470 ● www.e-iji.net

April 2022 ● Vol.15, No.2
p-ISSN: 1694-609X
pp. 119-138

Article submission code: 20210516224712

Received: 16/05/2021
Revision: 21/09/2021

Accepted: 13/10/2021
OnlineFirst: 09/01/2022

Impact of Formative Assessment Instructional Approach on Students' Mathematics Achievement and their Metacognitive Awareness

Ruth Nanjekho Wafubwa
Doctoral School of Education, Faculty of Humanities and Social Sciences, University of Szeged-Hungary, Hungary, *ruthnanje@gmail.com*

Csaba Csíkos
University of Szeged, MTA-SZTE Metacognition Research Group, Hungary, *csikoscs@edpsy.u-szeged.hu*

In the present study, the impact of formative assessment on secondary school students' mathematics achievement and their metacognitive awareness was examined. The participants included 164 (84 boys) grade 11 students and four mathematics teachers from four secondary schools in Kenya. A quasi-experimental pretest-posttest non-equivalent group design was used to determine if there was any significant difference in the students' mathematics achievement scores and their metacognitive awareness after a formative assessment intervention. The influence of gender on mathematics achievement and metacognitive awareness was also examined. Formative assessment was conceptualized as an instructional approach consisting of five strategies. After controlling for pretest scores, the results revealed that students who were taught using formative assessment instructional approach outperformed those taught using conventional approach for both mathematics achievement ($p = 0.014$, $\eta 2 = 0.38$) and metacognitive awareness ($p < 0.001$, $\eta 2 = 0.47$). Analysis based on gender did not reveal any gender influence on students' mathematics achievement ($p = 0.571$, $\eta 2 = 0.002$) and their metacognitive awareness ($p = 0.287$, $\eta 2 = 0.008$). The current study adds knowledge to the limited empirical evidence regarding the impact of formative assessment conceptualized as an instructional approach on students' mathematic's achievement and their metacognitive awareness.

Keywords: formative assessment, instructional approach, mathematics, students' achievement, metacognitive awareness

INTRODUCTION

Formative assessment is deemed to be a beneficial approach to instruction (Wiliam, 2011). The ground-breaking investigation on the impact of formative assessment (Black & Wiliam, 1998) has led to more research related to formative assessment. However,

Citation: Wafubwa, R. N., & Csíkos, C. (2022). Impact of formative assessment instructional approach on students' mathematics achievement and their metacognitive awareness. *international journal of instruction*, *15*(2), 119-138. https://doi.org/10.29333/iji.2022.1527a

various concerns regarding both the concept and efficacy of formative assessment have been raised by researchers (Bennett, 2011; Dann, 2014). Although inconsistencies in defining formative assessment and how it is applied have been notable concerns, increasing research on formative assessment is being conducted. This is reflected in the most recent review on how formative assessment strategies have been employed as well as their impact on students' learning outcomes (Wafubwa, 2020). According to the review, most scholars conceptualize formative assessment in terms of specific strategies but not formative assessment as a whole. Krijgsman et al. (2019) for instance focused on goal clarification and feedback, which constitute two strategies of formative assessment.

Studies on feedback have revealed the powerful impact that feedback has on achievement (Hattie & Timperly, 2007). After conceptualizing formative assessment as formative feedback, Pinger et al. (2018) found that feedback embedded in instruction can enhance students' performance. Although other studies have also explored the impact of formative feedback (e.g., Cutumisu & Schwartz, 2018; Kyaruzi et al., 2019), only a paucity of experimental studies on the influence of feedback has been carried out, particularly in secondary schools (Van der Kleij et al., 2015).

In most studies, formative assessment has been conceptualized as peer assessment (Vanderhovn et al., 2015; Hsia et al., 2016; Rotsaert et al., 2018; Tsivitanidou et al., 2018). In Black and Wiliam's framework, peer assessment strategy requires learners to be resourceful with each other in instructional processes. Peer assessment of given criteria can either be written or imparted verbally (Boud & Falchikov, 2007). According to Wiliam (2011), peer assessment is more productive when the focus is on improvement as opposed to evaluation. Empirical studies that have assessed the influence of peer assessment in different learning conditions have demonstrated that it enhances learning outcomes. Rotsaert et al. (2018), in their study involving peer assessment and feedback, showed that reciprocal peer assessment ensured immediate feedback. Tsivitanidoue et al. (2018) also utilized peer assessment as a learning tool.

Although there are a limited number of studies conceptualizing formative assessment as self-assessment, especially in the secondary schools' context, peer assessment is closely associated with self-assessment. Nikou and Economides (2016) studied the effect of self-assessment on students' motivation and achievement. In other studies, self-assessment has been conceptualized as part of self-regulation (e.g., Panadero et al., 2016; Panadero et al., 2017). These studies seem to suggest that the learning benefits of self-assessment can only be realized if students' are trained in self-regulation skills because it is innately difficult to acquire accurate self-knowledge (McDonald & Boud, 2003; Dunning et al., 2004).

Although most studies have examined teachers and students' perceptions, scholars have conceptualized formative assessment as a combination of five strategies (e.g., Burner, 2016; Dobish & Meyer, 2017; Saito & Inoi, 2017; Kippers et al., 2018; Ozan & Kincal, 2018; Johnson et al., 2019; Widiastuti et al., 2020). Few experimental studies related to the impact of formative assessment have demonstrated its positive influence on students' learning outcomes. Vogelzan and Admiraal (2017), who studied the impact of formative assessment found an improvement in students' chemistry performance. Pinger et al.

(2018), who conceptualized formative assessment as a tool embedded in the curriculum, revealed no improvement in the quality of instruction. Andersson and Palm (2017), after conceptualizing formative assessment as comprising different strategies, demonstrated improvement in students' achievement. Formative assessment has also been conceptualized in terms of tests and/or questions given to students at regular intervals to assess their learning (Heritage & Heritage, 2013). In the latter study, formative assessment was visualized as continuous assessment tests.

Self-assessment as a formative assessment strategy has been posited to enhance students' metacognitive awareness (Andrade, 2010; Taras, 2010) because students who participate in self-assessment monitor their thinking processes and are able to assess their learning process in general. Metacognition has been commonly described as the knowledge related to one's thought processes as well as the regulation of cognitive activities (Flavell, 1979). Thus, metacognition comprises two facets: metacognitive knowledge and metacognitive regulation. Schraw and Moshman (1995) noted that the acquisition of metacognitive awareness can be promoted by instructional strategies that activate students' self-knowledge and regulatory skills. Studies have shown that training students' metacognition influenced learning outcomes positively (Csíkos & Steklács, 2010; Roll et al., 2011; Naseri et al., 2017; Naful et al., 2021). One may conclude that self-assessment emphasizes high levels of metacognition, which influence learning styles and consequently learning achievement.

The present study

Formative assessment in the present study was employed as an instructional approach to enhance students' mathematics achievement on proportional reasoning skills and improve their metacognitive awareness. Black and Wiliam's (2009) framework, which envisions formative assessment as a classroom practice comprising five strategies, was employed as the theoretical framework of the study. The five strategies, which are subsequently described, are supposed to be utilized by the teacher, learner, and peer to identify and address learning gaps.

The first strategy may be described as sharing learning goals and criteria for attaining these goals. This strategy requires learners to know in which direction they are heading. Furthermore, it is important that the teacher involves learners in understanding what success looks like. The second strategy is effective classroom discussions. Discussions that primarily involve questioning are meant to reveal students' comprehension. Through discussion and questioning, the teacher can collect evidence of students' learning. The third strategy involves the provision of feedback that is not only given by the teacher but also by the learners and their peers. Through feedback, the learner is able to discover whether the learning goals are being attained. The teacher can also adjust instructional approaches to attend to students' needs. Peer assessment in which students act as each other's instructional resources is the fourth strategy. Peer assessment is beneficial because learners work in collaboration toward a common goal. The fifth strategy comprises self-assessment which involves activating students to own their learning. Self-assessment is not a stand-alone strategy but rather has to be incorporated in other formative assessment strategies (Wiliam, 2011). Black and Wiliam's (2009) framework postulates that these five strategies should be guided by three learning processes.

The first process in Black and Wiliam's framework involves identifying the direction in which the learners are heading, the second process encompasses establishing their current position, and the third process comprises knowing how they will reach their final destination. In accordance with this framework, the present study conceptualized formative assessment as an instructional approach that encompasses five strategies, three processes, and three agents, namely: teacher, student, and peer (Table 1).

Table 1
Features of formative assessment

	Where the learner is going	Where the learner is right now	How to get there
Teacher	1. Intentions for learning and criteria for success	2. Classroom discussion	3. Feedback
Peer	Understanding and sharing learning intentions and criteria for success	4. Peer assessment	
Learner	Understanding learning intentions and criteria for success	5. Self-assessment	

Adapted from Wiliam and Thompson (2008)

Although research on formative assessment and its impact on students' achievement has extensively been done, a few studies have focused on formative assessment as integration of five strategies. As noted previously, most studies have focused on specific strategies and in particular, on feedback and peer assessment. Furthermore, research has rarely examined specific skills in mathematics. However, this study is novel in that proportional reasoning skills in mathematics and students' metacognition were explored. Proportional reasoning is a crucial life skill utilized in day-to-day decision-making (Howe et al., 2011). Proportional reasoning in mathematics is among the areas in which students perform dismally in the Kenyan mathematics curriculum. It was considered that if students' proportional reasoning improved, their overall performance in mathematics would also improve and their metacognitive awareness would increase too. The present study thus contributes knowledge to the verifiable impact of formative assessment conceptualized as an instructional approach on students' mathematics achievement and their metacognitive awareness.

Research questions

In this study, the following four research questions were answered:

1. Is there a significant difference in the students' performance on mathematics posttest between the intervention and control groups?
2. Does gender influence students' performance on mathematics posttest? Is there a significant interaction between gender and the type of teaching approach, and mathematics posttest scores?
3. Is there a significant difference in the students' ratings on their levels of metacognitive awareness between the intervention and the control groups after the treatment?

4. Does gender influence students' metacognitive awareness after the treatment? Is there a significant interaction between gender and the type of teaching approach, and metacognitive awareness posttest scores?

METHOD

Participants

The participants included 164 grade 11 students (84 boys) from four low achieving rural secondary schools in the western part of Kenya. While two of the schools were randomly assigned to the experimental group, the remaining two formed the control group. There were 84 participants in the intervention group and 80 in the control group. Although the four schools had two classes each, only one class was selected randomly to take part in the study. Table 2 shows the demographics of the sample. Four teachers, one from each school voluntarily participated in this research. The participants had similar socio-economic and cultural backgrounds. In addition, the four schools were classified as sub-county schools according to Kenya's classification of schools. Therefore, one may deduce that the participants had similar characteristics in relation to their socio-economic background and academic performance.

Table 2
Sample demographics

Group	School/Class	Gender	N
Control	A	Boys	40
	B	Girls	40
Experimental	C	Boys	44
	D	Girls	40

Design

This study adopted a quasi-experimental pretest-posttest non-equivalent group design where schools were randomly assigned to either intervention group or control group. The intervention group consisted of 84 students whereas the control group comprised 80 students (Table 2). While the intervention schools were subjected to a formative assessment instructional approach, the control schools were taught by employing a conventional approach. The teachers taught the same content and matching tasks were given to the schools in both groups. Furthermore, both groups were given identical pretests and posttests on rates, ratios, and proportions. The participants also completed the Junior Metacognitive Awareness Inventory (Jr. MAI), which measured their levels of metacognitive awareness.

Procedure

This study comprised four phases (Figure 1). During the first phase, the four participating teachers were exposed to the proportional reasoning topic and the subtopics that they were expected to teach. Training sessions on formative assessment strategies and the implementation thereof were conducted with the teachers in the intervention group. On the other hand, teachers in the control group did not receive any training on formative assessment. In the second phase, students in both groups completed the pretest on the Proportional Reasoning Test (PRT) and their levels of metacognitive awareness were measured using the Jr. MAI. The teachers were given instructions on how to administer the test and questionnaire. Phase three was an implementation phase where teaching and learning took place in both conditions. Fourteen 40-minute lessons were conducted to complete the intervention in four weeks. While the teachers in the experimental group utilized the materials and tasks given during the training session, those in the control group employed a conventional approach when teaching. The five strategies of formative assessment (Table 1) were implemented in the 14 lessons in the intervention group. In the final phase, participating students did the PRT and completed the Jr. MAI.

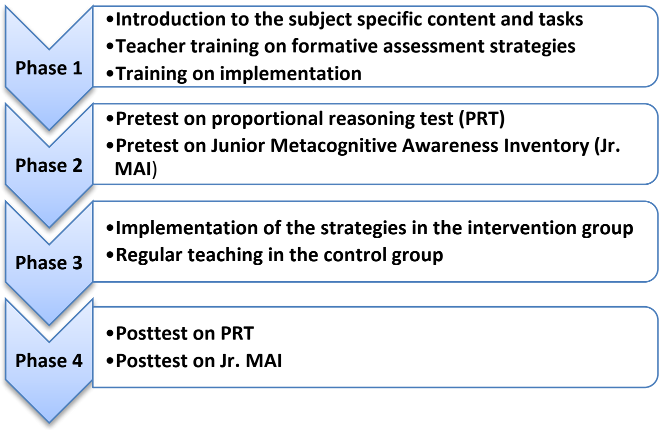


Figure 1
Intervention phases

Teacher training and experimental conditions

The four participating teachers had already been informed about the project during the survey study a year previously and had decided to participate willingly. The training, which comprised workshops, was divided into two parts. During the first, which lasted a day, proportional reasoning and specific areas that had to be covered under each topic were introduced. The teachers also discussed how they could handle the five areas of proportional reasoning in 14 lessons over a period of four weeks. They agreed on a flexible four-week scheme that could be employed to guide them teach the topic. The time allocation was in line with the recommended time stipulated in the Kenyan mathematics curriculum (KICD, 2017). The second part of the training was conducted over two days and involved only the teachers in the intervention group who received training on how to implement formative assessment strategies.

The teacher training in the intervention group was centered on five strategies of formative assessment as well as how the strategies could be implemented. During the training, implementation challenges were addressed, and the teachers agreed on a three-step flexible guideline to implement the strategies (Appendix A1). During the training session, teachers were also provided with supplementary materials on formative assessment strategies. The training program and specific areas that the training focused on are displayed in Appendix A2. In the experimental classes, the teacher first stated the rationale for each particular lesson and briefly engaged students in designing their success through questioning and discussion. Second, the teacher gave tasks to students individually which was followed by a group discussion. The discussion enabled students to obtain feedback from each other. Finally, the teacher gave feedback to individual groups in order to generate and share ideas in groups. These ideas were subsequently discussed with the whole class. The groups, which were also referred to as study groups, comprised four to five students with mixed abilities. The whole process involved the students and teachers jointly identifying and communicating the learning and performance goals. The participants' current levels of understanding were assessed and strategies and skills to reach goals were generated.

Measures

Proportional Reasoning Test (PRT)

Previously, the researchers and two mathematics teachers who had considerable experience in teaching high school mathematics and were also national examiners developed a written mathematics test on proportional reasoning skills. The 10 items of PRT cover all the content areas on rates, ratios, and proportions, a topic covered in the Kenya secondary schools mathematics curriculum. The test was constructed from word problems relating to real-life situations and examines five aspects of proportional reasoning: missing values, associated sets, mixtures and proportions, comparison problems, and stretcher (Appendix B). Two items assess each aspect and there is an equal distribution of marks across the 10 items.

The content validity was determined by a team of mathematics subject experts. Furthermore, item-level analysis was performed by examining the difficulty and discrimination indices. The PRT test was piloted on a sample of 45 students and analysis of items showed the difficulty level ranged between 0.39 and 0.61, thus implying a moderately difficult test. All the items have a higher cognitive demand and therefore, require students to employ self-regulatory skills to solve them. Therefore, formative assessment which involves self-regulation strategies was deemed to be the best approach to enhance students' proportional reasoning. For the present sample, Cronbach's alpha coefficient was 0.72, suggesting acceptable reliability. More details on how the PRT was developed and validated can be found in Wafubwa et al. (2020).

Junior Metacognitive Awareness Inventory (Jr. MAI)

This inventory was constructed by Sperling et al. (2002) to assess young adults' metacognition as well as a tool to gauge classroom interventions. In this study, the adapted Jr. MAI was employed to measure students' level of metacognitive awareness before and after a formative assessment intervention. The inventory comprised 18 items, which measured the knowledge and regulation dimensions of metacognition. Examples of items in the knowledge dimension were: "*I know when I understand something*" and

"*I can make myself learn when I need to*". On the regulation dimension, examples of items were: "*I think about what I need to learn before I start working*" and "*I pay attention to important information*" (Sperling et al., 2002 p.76). Items were evaluated on a 5 point scale ranging from 1 (not at all) to 5 (always).

The pretest and posttest questionnaires were not completed by five and nine participants, respectively. Thus, the data of the 14 incomplete questionnaires were excluded from the analysis. The two-factor solution of principal component factor analysis was performed to ascertain validity. The Kaiser-Meyer-Olkin (KMO) was 0.652 and Bartlett's test value was significant (χ^2 (351) = 940.316, $p < .001$). While 10 factors loaded on the knowledge dimension, eight factors loaded on the regulation dimension. Although some of the factors intended for the knowledge dimension following the original scale loaded on the regulation dimension and vice versa, this was not a concern in our current study because the intention was to measure students' metacognition in general. Cronbach's alpha coefficient revealed that the reliability for the whole scale was 0.78, which is considered to be acceptable internal consistency.

Study Variables

There were two active independent variables in this study. Formative assessment was employed as a between-groups independent variable (IV) with two levels: formative assessment and no formative assessment. The second active IV was, change over time, which was a within-subjects IV with two levels: pretest and posttest. Gender was utilized as an attribute IV with two levels: male and female. Pretest scores on both achievement and metacognitive awareness were employed as a covariate in the Analysis of covariance (ANCOVA). Two dependent variables were measured in this study: students' mathematics achievement, which was measured using PRT, and metacognition, which was assessed by employing the Jr. MAI.

The implementation process

The pretest was administered to both the experimental and control classes a week after the teacher training workshops. The teachers in the experimental group implemented formative assessment through a process of reciprocal classroom interaction that involved the teachers, students, and learning resources. The implementation involved a three-step guide that specified strategies to be used during each step (Appendix A1). In the first step, strategies that described learning intentions and success criteria were employed. In the second step, self-assessment, peer assessment, and discussion strategies were utilized. The third step involved peer assessment, self-assessment, and feedback strategies. The guide included all five strategies blended in the lessons.

In the course of the implementation process, the lead researcher conducted a follow-up twice a week to ensure the implementation was being conducted as planned. However, it was not possible to observe the teaching because of COVID 19 pandemic-related restrictions. The posttest was administered one month after the pretest. During the administration of both the pretest and posttest, the teachers were given a set of instructions to ensure test fidelity. The instructions included the time required to complete the test, authorized instruments, and spacing of students when doing the test.

The scoring of the tests was conducted externally by the lead researcher and other two experienced teachers from a different region to that of the study. Scoring was blinded in that the examiners were unaware of which group the students belonged. Once the scoring was complete, the teachers were given back the students' scripts so as to provide feedback to inform instruction.

Data analysis

In the study, students were the unit of analysis. Both the pre and posttest data obtained using the two research instruments were examined for parametric tests assumptions. While the Shapiro-Wilk test was employed to examine whether the data were normally distributed, Levene's test examined homogeneity of variance. Analysis of covariance (ANCOVA) was utilized in testing for significant variations in the posttest scores on achievement and metacognition, thus assessing the impact of formative assessment. The effect sizes were determined by partial Eta Squared values.

FINDINGS

Mathematics achievement

The impact of formative assessment was assessed through ANCOVA. The group variable was formative assessment. Whereas the independent variable was the students' pretest mean score, the dependent variable was the students' posttest mean score. The results of the Shapiro-Wilk test for both groups were not significant ($p > 0.05$), thus implying a normal distribution for the covariate and dependent variable. Furthermore, an examination of boxplots did not reveal any extreme outliers. Levene's test found non-significant results ($p = 0.200$) for equality of error variance. The homogeneity of regression slopes assumption was also examined and the effect was found to be non-significant ($p = 0.335$) hence, the assumption was met.

After using pretest scores as the covariate, the ANCOVA results showed a notable variation in the posttest scores between the students in the intervention group and the control group, $F(1, 159) = 6.227$, $p = 0.014$, $\eta2 = 0.38$. The results imply that students who were taught by employing formative assessment strategies improved in their achievement on PRT in comparison to those who were taught by using conventional approaches. The adjusted means, standard errors (SE), means (M), and standard deviations (SD) for the posttest groups are presented in Table 3.

Table 3
Posttest mean scores

Group	N	Adjusted M	SE	M	SD
Experimental	84	1.69	0.056	1.70	0.58
Control	80	1.49	0.057	1.48	0.52

A two-way ANCOVA was conducted to show how gender influenced the posttest scores. The assumptions of ANCOVA were all met. Both the Shapiro-Wilk test and Levene's test were non-significant. The results showed that gender had no significant influence on the posttest scores, $F(1, 159) = 0.322$, $p = 0.571$, $\eta2 = 0.002$. Furthermore, no significant interaction between gender and the type of teaching approach, and the mathematics posttest scores was exhibited, $F(1, 159) = 0.347$, $p = 0.557$, $\eta2 = 0.002$.

The adjusted means, SE, means (M) and SD for the posttest mean scores based on gender are displayed in Table 4. The pretest mean scores were used as the covariate.

Table 4
Posttest mean scores based on gender

Gender	Group	N	Adjusted M	SE	M	SD
Boys	Control	40	1.56	0.072	1.53	0.47
	Experimental	44	1.68	0.066	1.70	0.57
Girls	Control	40	1.45	0.073	1.42	0.56
	Experimental	40	1.66	0.064	1.71	0.60

Metacognition

The influence of formative assessment on students' metacognitive awareness was determined by conducting ANCOVA. After checking all the assumptions of ANCOVA, the results of the analysis indicated a notable difference in the metacognition posttest scores between the formative assessment and control groups, $F(1, 145)= 128.260$, $p < 0.001$, $\eta 2 = 0.469$. The ANCOVA results revealed that metacognitive awareness rating was higher among the students who received instruction using formative assessment than those who were taught using conventional methods. In Table 5, the adjusted means, SE, means (M), and SD for the posttest scores while using the pretest mean scores as the covariate are shown.

Table 5
Posttest mean scores

Group	N	Adjusted M	SE	M	SD
Intervention	77	4.30	0.26	4.26	0.29
Control	73	3.88	0.26	3.91	0.49

A two-way ANCOVA was used to estimate the effect of gender on the posttest scores of metacognition. The results showed no significant effect of gender on the posttest scores, $F(1, 145)= 1.142$, $p = 0.287$, $\eta 2 = 0.008$. In addition, no significant interaction between gender and the type of teaching approaches on the posttest scores was exhibited, $F(1, 145) = 0.088$, $p = 0.767$. $\eta 2 = 0.001$. These results suggest that both male and female students responded in a similar way to the teaching method. The adjusted means, SE, means (M), and SD for the posttest mean scores based on gender with pretest mean scores as the covariate are displayed in Table 6.

Table 6
Posttest mean scores based on gender

Gender	Group	N	Adjusted M	SE	M	SD
Boys	Control	35	3.89	0.038	3.98	0.46
	Experimental	41	4.32	0.035	4.27	0.28
Girls	Control	38	3.86	0.037	3.85	0.51
	Experimental	36	4.27	0.038	4.25	0.31

DISCUSSION

In this research, the influence of formative assessment on students' mathematics achievement and their metacognitive awareness was explored. Four research questions guided the study. Research question one sought to establish if there was a notable difference in the posttest scores between intervention and control groups after the treatment. Results exhibited a significant difference between the posttest scores of the two groups with a medium effect size ($\eta 2 = 0.38$) after controlling for the pretest scores. This implies that students who were exposed to formative assessment strategies performed better than those who were taught conventionally. Other studies have also demonstrated that the utilization of formative assessment strategies improves students' performance (Ozan & Kıncal, 2018; Pinger et al., 2018; Vogelzan & Admiraal, 2017). However, most of these studies focused on one or two strategies and not on formative assessment as a combination of strategies. Formative assessment in the current study was conceptualized as an instructional approach encompassing five strategies (Figure 1), which are embedded in instruction. Anderson and Palm (2017) who also conceptualized formative assessment as a combination of strategies found a significant effect on students' achievement (Cohen's $d = 0.66$). The effect size they found was larger than that of this study. While we employed students as the unit of analysis, Anderson and Palm used teachers as the unit of analysis. This is possibly the reason for the difference in the effect sizes.

However, some considerations should be taken when interpreting the effect sizes related to formative assessment. The first consideration is on how formative assessment is conceptualized and the second is how it is implemented (Bennett, 2011). Research has indicated that the efficacy of formative assessment has been hampered by poor implementation processes (Randel et al., 2016). Therefore, teachers' preparation and support are crucial for the execution of formative assessment. Although the duration for teacher training for this study was short, professional development may not be the sole determining factor for the successful implementation of an intervention (Johnson et al., 2019; Randel et al., 2016; Yin & Buck, 2019). Studies have further shown that apart from inadequate professional development, teachers fail to implement formative assessment because of their heavy workload and lack of motivation (Crichton & McDaid, 2016; Jacoby et al., 2014). Because the teachers in this study received external support and participated willingly in the study, we can deduce that the improvement in the experimental group was due to formative assessment strategies.

The second research question was concerned with whether gender influenced mathematics achievement scores after the intervention and whether a notable interaction between gender and the type of teaching approach and the posttest scores were exhibited. The results revealed that after controlling for pretest scores, gender did not influence the posttest scores. There was also no evident interplay between gender and the type of teaching approaches, and the posttest scores. The results suggest that formative assessment had a similar influence on the learning of both male and female students. Therefore, one may infer that the improvement in achievement in the intervention group was associated with the formative assessment instructional approach and gender did not influence this approach. Recent studies on gender and achievement have also shown that gender does not influence mathematics achievement (Lindberg et al., 2010; Louis & Mistele, 2012; Scheiber et al., 2015).

The third research question focused on whether the teaching approach had an influence on the students' posttest scores on metacognitive awareness. The results revealed that students who received instruction using formative assessment had a higher metacognitive rating than those who were taught using conventional methods. The items on Jr. MAI assessed student's metacognitive awareness, which comprised knowledge and skills dimensions. Knowledge of cognition involves awareness of and knowledge about one's cognition (Harris et al., 2010). On the other hand, metacognitive skills involve planning, monitoring, and evaluating learning processes (Veenman & Beishuizen, 2004). Although our literature search did not yield studies related to the influence of formative assessment on students' metacognition, scholars have demonstrated that formative assessment and metacognitive skills are related (Baas et al., 2014; Wafubwa & Csíkos, 2021). Empirical studies have also shown the benefits of metacognition on students' achievement, especially when students are trained to be metacognitive (Csíkos & Steklács, 2010; Roll et al., 2011; Naseri et al., 2017; Dafik & Rohim, 2019; Naful et al., 2021).

Being metacognitive implies that one is conscious of his or her thought processes and can regulate cognition through processes such as monitoring, planning, and evaluating. Formative assessment strategies, in particular, self-assessment and feedback, also involve the self-regulation processes of monitoring, planning, and evaluating. Therefore, it is possible that the formative assessment strategies improved students' metacognition, which is reflected in the higher ratings on the posttest scores. However, although we acknowledge that it is not always possible for students to have a true knowledge of themselves, based on their performance on the PRT, one may deduce that the ratings on metacognitive awareness inventory could be a reflection of what students feel about their metacognition. However, proportional reasoning is only one skill in mathematics. Therefore, the results should be elucidated within proportional reasoning skills' context but not generalized to other areas in mathematics.

The fourth research question was concerned with the influence of gender on metacognitive awareness posttest scores and whether there was an interaction between gender and the type of teaching approach and metacognitive awareness posttest scores. The results revealed that gender did not have a significant influence on metacognitive awareness posttest scores. Furthermore, no significant interaction between gender and teaching approach, and the posttest scores were found. Although research on gender and students metacognition has not been extensively studied, some studies have suggested that students' metagonition is not dependent on gender (Al Shabibi & Alkharusi, 2018; Çakici, 2018; Siswati & Corebima, 2017).

CONCLUSION

The findings of the present study indicate that formative assessment strategies based on students' needs can lead to improved learning outcomes when employed. The formative assessment approach used in this study was planned following the challenges that students encountered with proportional reasoning in mathematics. Students were thus not able to solve mathematics questions that needed the use of metacognitive strategies such as planning, monitoring, and evaluating. However, it cannot be assumed that the same results may be realized with other mathematics topics. Rather, it is dependent on how teachers conceptualize formative assessment. When formative assessment is conceptualized as continuous assessment tests, all of the five strategies may not be employed, which may result in an insignificant impact. On the contrary, if teachers view formative assessment as a classroom practice that can show evidence of student learning and enable them to make decisions on how to improve instruction, all of the five strategies will be utilized for better learning outcomes. This study has revealed that formative assessment strategies can improve the performance of low achieving students and also improve their metacognitive awareness. The results of this study can benefit teachers and curriculum developers in designing formative assessment intervention programs that can boost students' achievement in mathematics and improve their metacognitive awareness. It is recommended that future studies explore other topics in mathematics using the same formative assessment approaches.

ACKNOWLEDGMENT

This work was supported by the Research Programme for Public Education Development of the Hungarian Academy of Sciences.

REFERENCES

Al Shabibi, A. & Alkharusi, H. (2018). Mathematical problem-solving and metacognitive skills of 5th grade students as a function of gender and level of academic achievement. *Cypriot Journal of Educational Science*, *13*(2), 385−395.

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. https://doi.org/10.1016/j.learninstruc.2016.12.006.

Andrade, H. (2010). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In H. J. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 90–105). Routledge.

Baas, D., Castelijns, J., Vermeulen, M., Martens, R., & Segers, M. (2014). The relation between Assessment for Learning and elementary students' cognitive and metacognitive strategy use, *British Journal of Educational Psychology*, *85*(1), 36–46. https://doi.org/10.1111/bjep.12058.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, *5*(1), 7-74.https://doi.org/10.1080/0969595980050102.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation, and Accountability (formerly: Journal of Personnel Evaluation in Education)*, *21*(1), 5-31.DOI:10.1007/s11092-008-9068-5

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in education: principles, policy & practice*, *18*(1), 5-25. http://dx.doi.org/10.10800969594X.2010.513678.

Boud, D., & Falchikov, N. (Eds.). (2007). *Rethinking assessment in higher education: Learning for the longer term.* Routledge.

Burner, T. (2016). Formative assessment of writing in English as a foreign language. *Scandinavian Journal of Educational Research*, *60*(6), 626-648

https://doi.org/10.1080/00313831.2015.1066430

Crichton, H., & McDaid, A. (2016). Learning intentions and success criteria: learners' and teachers' views. *The Curriculum Journal*, *27*(2), 190-203. https://doi.org/10.1080/09585176.2015.1103278

Csíkos, C., & Steklács, J. (2010). Metacognition-based reading intervention programs among fourth-grade Hungarian students. In *Trends and Prospects in Metacognition Research* (pp. 345-366). https://doi.org/10.1007/978-1-4419-6546-216.

Cutumisu, M., & Schwartz, D. L. (2018). The impact of critical feedback choice on students' revision, performance, learning, and memory. *Computers in Human Behavior, 78*, 351-367. https://doi.org/10.1016/j.chb.2017.06.029

Çakici, D. (2018). Metacognitive Awareness and Critical Thinking Abilities of Pre-Service EFL Teachers. *Journal of Education and Learning*, *7*(5), 116-129. https://doi.org/10.5539/jel.v7n5p116

Dafik, S. B., Irvan, M., & Rohim, M. A. (2019). The Analysis of Student Metacognition Skill in Solving Rainbow Connection Problem under the Implementationof Reseach-Based Learning. *International Journal of Instruction*, *12,* 593-610.

Dann, R. (2014). Assessment as learning: blurring the boundaries of assessment and learning for theory, policy, and practice. *Assessment in Education: Principles, Policy & Practice*, *21*(2), 149-166. https://doi.org/10.1080/0969594X.2014.898128

Dobish, M., Griffiths, J., & Meyer, R. (2017). Improving Teaching and Learning Using the Keeping Learning on Track Professional Development Program and Strategies. *Journal for Leadership and Instruction*, *16*(1), 22-26.

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological science in the public interest*, *5*(3), 69-106. https://doi.org/10.1111/j.1529-1006.2004.00018.x

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive development inquiry. *American psychologist, 34*(10), 906.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, *77*(1), 81-11. https://doi.org/10.3102/003465430298487

Harris, K. R., Santangelo, T., & Graham, S. (2010). Metacognition and Strategies instruction in Writing. In H. S. Schneider and W. Waters, *Metacognition, Strategy Use, and Instruction (*pp. 226-256). The Guilford Press

Heritage, M., & Heritage, J. (2013). Teacher questioning: The epicenter of instruction and assessment. Applied Measurement in Education, *26*(3), 176-190. https:// dx.doi.org/10.1080/08957347.2013.793190

Howe, C., Nunes, T., & Bryant, P. (2011). The rational number and proportional reasoning: Using intensive quantities to promote achievement in mathematics and science. *International Journal of Science and Mathematics Education*, *9* (2), 391-417. https://doi.org/10.1007/s10763-010-9249-9

Hsia, L. H., Huang, I., & Hwang, G. J. (2016). A web-based peer-assessment approach to improving junior high school students' performance, self-efficacy, and motivation in performing arts courses. *British Journal of Educational Technology*, *47*(4), 618-632. https://doi.org/10.1111/bjet.12248

Jacoby, J., Heugh, S., Bax, C., & Branford- White, C. (2014). Enhancing learning through formative assessment. *Innovations in education and teaching international. 51*(1). pp. 72-83.

Johnson, C. C., Sondergeld, T. A., & Walton, J. B. (2019). A Study of the Implementation of Formative Assessment in Three Large Urban Districts. *American Educational Research Journal*, *56*(6), 2408-2438. https://doi.org/10.3102/0002831219842347

Kenya Institute of Curriculum Development, KICD. (2017). Republic of Kenya: Basic Education Curriculum Framework. Republic of Kenya, KICD.

Kippers, W. B., Wolterinck, C. H., Schildkamp, K., Poortman, C. L., & Visscher, A. J. (2018). Teachers' views on the use of assessment for learning and data-based decision making in classroom practice. *Teaching and Teacher Education*, *75*, 199-213. https://doi.org/10.1016/j.tate.2018.06.015

Krijgsman, C., Mainhard, T., van Tartwijk, J., Borghouts, L., Vansteenkiste, M., Aelterman, N., & Haerens, L. (2019). Where to go and how to get there: Goal clarification, process feedback and students’ need satisfaction and frustration from lesson to lesson. *Learning and Instruction*, (*61),* 1-11. https://doi.org/10.1016/j.learninstruc.2018.12.005

Kyaruzi, F., Strijbos, J. W., Ufer, S., & Brown, G. T. (2019). Students’ formative assessment perceptions, feedback use and mathematics performance in secondary schools in Tanzania. *Assessment in Education: Principles, Policy & Practice*, *26*(3), 278-302.

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin, 136*(6), 1123–1135. https://doi.org/10.1037/a0021276

Louis, R. A., & Mistele, J. M. (2012). The differences in scores and self-efficacy by student gender in mathematics and science. *International Journal of Science and Mathematics Education*, *10*(5), 1163-1190. https://doi.org/10.1007/s10763-011-9325-9

McDonald, B., & Boud, D. (2003). The impact of self-assessment on achievement: The effects of self-assessment training on performance in external examinations. *Assessment in education: principles, policy & practice*, *10*(2), 209-220. https://doi.org/10.1080/0969594032000121289

Naufal, M. A., Abdullah, A. H., Osman, S., Abu, M. S., & Ihsan, H. (2021). The effectiveness of infusion of metacognition in van Hiele model on secondary school students' geometry thinking level. *International Journal of Instruction*, *14*(3), 535-546. https://doi.org/10.29333/iji.2021.14331a

Naseri, M., Shoae, K. M., Effati, M, M. (2017). The effectiveness of metacognitive skills training on increasing academic achievement. *Iranian journal of educational sociology,* 1(3), 83-88

Nikou, S. A., & Economides, A. A. (2016). The impact of paper-based, computer-based and mobile-based self-assessment on students' science motivation and achievement. *Computers in Human Behavior*, *55*, 1241-1248. https://doi.org/10.1016/j.chb.2015.09.025

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. http://dx.doi.org/10.12738/estp.2018.1.0216

Panadero, E., Brown, G. T., & Strijbos, J. W. (2016). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review*, *28*(4), 803-830. https://doi.org/10.1007/s10648-015-9350-2.

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, *22*, 74-98. https://doi.org/10.1016/j.edurev.2017.08.004

Pinger, P., Rakoczy, K., Besser, M., & Klieme, E. (2018). Implementation of formative assessment–effects of quality of programme delivery on students' mathematics achievement and interest. *Assessment in Education: Principles, Policy & Practice*, *25*(2), 160-182.https://doi.org/10.1080/0969594X.2016.1170665

Randel, B., Apthorp, H., Beesley, A., Clark, T., & Wang, X. (2016). Impacts of professional development in classroom assessment on teacher and student outcomes. The Journal of Educational Research, 109(5), 491-502. http://dx.doi.org/10.1080/00220671.2014.992581

Roll, I., Aleven, V., McLaren, B. M., & Koedinger, K. R. (2011). Improving students' help-seeking skills using metacognitive feedback in an intelligent tutoring system. *Learning and instruction*, *21*(2), 267-280.

Rotsaert, T., Panadero, E., Estrada, E., & Schellens, T. (2017). How do students perceive the educational value of peer assessment in relation to its social nature? A survey study in Flanders. *Studies in Educational Evaluation*, (*53)*, 29-40. https://doi.org/10.1016/j.stueduc.2017.02.003

Saito, H., & Inoi, S. I. (2017). Junior and senior high school EFL teachers' use of formative assessment: A mixed-methods study. *Language Assessment Quarterly*, *14*(3), 213-233. https://doi.org/10.1080/15434303.2017.1351975

Scheiber, C., Reynolds, M. R., Hajovsky, D. B., & Kaufman, A. S. (2015). Gender differences in achievement in a large, nationally representative sample of children and adolescents. *Psychology in the Schools*, *52*(4), 335-348.

Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational psychology review*, *7*(4), 351-371. https://doi.org/10.1007/BF02212307

Siswati, B. H., & Corebima, A. D. (2017). The effect of education level and gender on students' metacognitive skills in Malang, Indonesia. *Advances in Social Sciences Research Journal*, *4*(4). DoI:10.14738/assrj.44.2813

Sperling, R. A, Howard, B. C, Miller, L. A, & Murphy, C. (2002). Measures of children's knowledge and regulation of cognition. *Contemporary Educational Psychology*, *27* (1), 51-79.

Taras, M. (2010). Student self-assessment: Processes and consequences. *Teaching in higher education*, *15*(2), 199-209.https://doi.org/10.1080/13562511003620027.

Tsivitanidou, O. E., Constantinou, C. P., Labudde, P., Rönnebeck, S., & Ropohl, M. (2018). Reciprocal peer assessment as a learning tool for secondary school students in modeling-based learning. *European Journal of Psychology of Education*, *33*(1), 51-73.

Vanderhoven, E., Raes, A., Montrieux, H., Rotsaert, T., & Schellens, T. (2015). What if pupils can assess their peers anonymously? A quasi-experimental study. *Computers & Education*, *81*, 123-132. https://doi.org/10.1016/j.compedu.2014.10.001

Van der Kleij, F. M., Feskens, R. C., & Eggen, T. J. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis. *Review of educational research*, *85*(4), 475-511.https://doi.org/10.3102/0034654314564881.

Veenman, M. V., & Beishuizen, J. J (2004). Intellectual and Metacognitive Skills of Novices While Studying Texts under Conditions of Text Difficulty and Time Constraint. *Learning and Instruction 14* (6), 621–640. https://doi.org/10.1016/j.learninstruc.2004.09.004

Vogelzang, J., & Admiraal, W. F. (2017). Classroom action research on formative assessment in a context-based chemistry course. *Educational Action Research*, *25*(1), 155-166. https://doi.org/10.1080/09650792.2016.1177564

Wafubwa, R. N. (2020). Role of Formative Assessment in Improving Students' Motivation, Engagement, and Achievement: A Systematic Review of Literature. *The International Journal of Assessment and Evaluation, 28* (1), 17-31.

Wafubwa, R. N., & Csíkos, C. (2021). Formative Assessment as a Predictor of Mathematics Teachers' Levels of Metacognitive Regulation. International Journal of Instruction, *14*(1), 983-998. https://doi.org/10.29333/iji.2021.14158a

Wafubwa, R. N., Opoku-Sarkodie, R., & Csíkos, C. (2020). Construction and Validation of a Test for Measuring Students' Proportional Reasoning on Rates, Ratios, and Proportions. *International Journal of Pedagogy, Policy, and ICT in Education*, (*8), 53-76.

Widiastuti, I. A. M. S., Mukminatien, N., Prayogo, J. A., & Irawati, E. (2020). Dissonances between Teachers' Beliefs and Practices of Formative Assessment in EFL Classes. *International Journal of Instruction*, *13*(1), 71-84. 71-84. https://doi.org/10.29333/iji.2020.1315a

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

Wiliam, D, Thompson, M. (2008) Integrating Assessment with Learning: What Will It Take to Make It Work? In: Dwyer, C. A, (ed.). *The Future of Assessment: Shaping Teaching and Learning.* (pp. 53-82). Routledge https://doi.org/10.4324/9781315086545

Yin, X., & Buck, G. A. (2019). Using a collaborative action research approach to negotiate an understanding of formative assessment in an era of accountability testing. *Teaching and Teacher Education*, (*80),* 27-38. https://doi.org/10.1016/j.tate.2018.12.018.

APPENDIX A1

Formative assessment Mathematics lesson plan guide for teachers

Stage	Strategy	Teaching action	Purpose
Step 1 The direction in which the student is moving (10 min)	Learning targets and standards for success	Present the problem to the students. Class discussion Clarify learning intentions	To elicit ideas of possible ways to solve the problem.
Step 2 Where the student is currently (15 minutes)	Discussions Peer and self-assessment	Use questions to enable students to determine where they are. Provide hints.	Diagnose students' strengths and limitations. Students' reflection
Step 3 How to get there (15 minutes)	Self and peer assessment Teacher's feedback	Explain and discuss solutions to the problems. Discuss special features.	Illustrate different strategies and their application in authentic situations.

APPENDIX A2

Formative assessment training schedule

Day 1	Session 1	• Introduction to formative assessment • Different concepts of formative assessment • Teachers' conceptions and experiences of formative assessment
	Session 2	• A conceptual framework for the study (Black & Wiliam, 2009): five strategies, three agents, and three processes
Day 2	Session 1	• A detailed discussion of each strategy and how each can promote learning (based on research reports) • Discussion on how the strategies can enhance metacognition (research evidence) • Teachers' views on different strategies
	Session 2	• Implementation of the strategies • Implementation challenges and how they can be addressed • Lesson plan guide

APPENDIX B

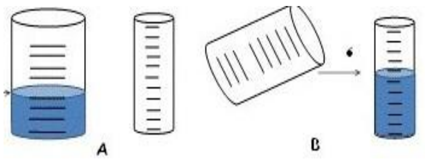
Proportional Reasoning Test (30 marks)

1. Last week, Mary answered 24 out of 30 questions correctly in an exam. This week, she answered 20 out of 24 questions correctly in another exam. For which exam did Mary have better results? Explain your answer.

2. Nafula bought 3 lollypops at 12 shillings and Atieno bought five lollypops for 20 shillings. Who used less amount of money? Explain your answer.

3. How many glasses of orange juice can you make with 12 cups of water if eight 8 cups of water can produce 14 glasses of orange juice? Show your calculations.

4. The diagrams below show two tins of different sizes but marked with the same scale on each of them. Oil is poured into the broad tin until it reaches the fourth mark. When the same oil is poured into the small tin as demonstrated in diagram B, it rises to the sixth mark. If both tins are emptied and oil is poured into the broad tin until it reaches the sixth mark, to what level cans this oil rise if it is poured into the small cylinder?



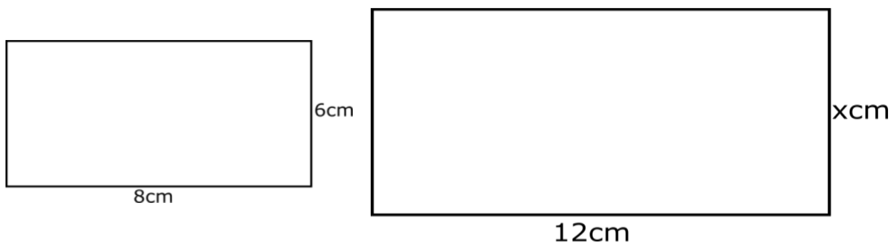
5. A group of 7 girls shares 3 chapatis equally and another group of 9 boys shares 4 chapatis equally. Which group gets a larger piece of chapatti? Why?

6. Mary has the option of working in Mombasa or Nairobi. She discovered that the workers in Mombasa work 8 hours per day and receive Ksh 24 000 every 15 days while those in Nairobi work 6 hours per day and receive Ksh 20 000 every 12 days. If Mary decides to work for 20 days, which job option will be best for her? Explain your answer.

7. Your father decides to give a piece of land as an inheritance to your three brothers, Joe, Alex, and Peter in the ratio 4:5:3. Peter being the firstborn feels he has already accumulated enough wealth and therefore, decides to share his portion equally with Joe and Alex. Calculate the ratio of Joe's share to Alex's share. Show your calculations.

8. In a mixture of 60 liters, the ratio of orange concentrate to water is 7:5. If the principal of a school wants to make orange juice for the students by using the ratio of 3:2, how many liters of water should he add to the mixture? Show your calculations.

9. The figures below show two similar rectangles. The height of the first rectangle is 6cm and the width is 8cm. The width of the second rectangle is 12cm. Explain how you would find the height of the larger rectangle.



10. The heights of two trees taken three years ago were eight feet for the tree (I) and ten feet for the tree (II). When the heights were taken today, tree (I) was 14 feet and tree (II) was 16 feet. Which of the trees increased most over the past three years?