

The Effect of Formative Testing and Self-Directed Learning on Mathematics Learning Outcomes

Mohamad Syarif SUMANTRI *

Jakarta State University, Indonesia

Retni SATRIANI

Jakarta State University, Indonesia

Received: 8 December 2015 / Revised: 28 January 2016 / Accepted: 7 February 2016

Abstract

The purpose of this research was to determine the effect of formative testing and self-directed learning on mathematics learning outcomes. The research was conducted at an elementary school in central Jakarta during the 2014/2015 school year. Seventy-two fourth-grade students who were selected using random sampling participated in this study. Data were obtained through testing and were analyzed using a two-line analysis of variance (ANOVA) according to the treatment design and level of self-directed learning. The results showed that (1) mathematics learning outcomes differ between students who are given formative essay tests and those who are given formative multiple choice tests; (2) there is an interaction effect between formative testing and self-directed learning on mathematics learning outcomes; (3) students with high levels of self-directed learning have better learning outcomes when given formative essay tests than when given formative multiple choice tests; and (4) students with low levels of self-directed learning show no difference in mathematics learning outcomes based on whether they are given formative essay tests or multiple choice tests.

Keywords: Formative test, Self-directed learning, Mathematics outcomes.

Introduction

The *Trends in International Mathematics and Science Study* (Mullis, Martin, Foy, & Arora, 2011) conducted in 2011 indicated that Indonesia's students have only a basic knowledge of mathematics, which is not sufficient to solve routine questions and problems or to manipulate mathematic forms and apply the necessary logical strategies. Indonesian students rank 38 out of 42 countries in mathematical achievement, with an average score of 386 based on the TIMSS scales (which range between 300 and 700). Other evidence is available from the results of *The Program for International Students Assessment* (PISA), which reported in 2012 that Indonesian students scored 375 for mathematical capability.

* ✉ Mohamad Syarif SUMANTRI, Jakarta State University, Indonesia Department of Elementary Education Post Graduate Studies Program (PPs), Universitas Negeri Jakarta Jalan Rawamangun Muka, Jakarta, 13220. Phone: +6221- 47213040, Email: dikdas.pps@unj.ac.id

Other student achievement indicators include the national examination, which is conducted every year and has resulted in scores below 7, which are unsatisfactory. The average mathematics scores on the national examination (scored on a scale of 0-10) for public elementary schools in Jakarta are shown in Table 1 below.

Table 1. *The Average National Examination Mathematics Scores for Public Elementary Schools in DKI Jakarta and Kepulauan Seribu for 2011/2012 and 2012/2013*

Region/City	Academic year	
	2011/2012	2012/2013
Central Jakarta	7.04	6.73
North Jakarta	6.77	6.64
East Jakarta	7.33	7.34
West Jakarta	6.57	6.55
South Jakarta	7.29	7.13
Kepulauan Seribu	6.24	6.32
Average	6.87	6.78

Source: www.simdik.info/hasilun/index.aspx.

The data above show that mathematics achievement is relatively low. Students struggle with mathematics for several reasons, including the abstract nature of mathematical figures, the fact that most students view mathematics as a frightening subject, students' difficulty in comprehending mathematics or lack of interest in studying mathematics, and students' disabilities related to mathematics learning skills. The National Council for Accreditation of Teacher Education (2003) stated that mathematics teachers should "possess a deep understanding of how students learn mathematics and use the pedagogical knowledge specific to mathematics teaching and learning" (as cited in Osborne, 2015, p. 23).

Osborne (2015) concluded that educators must embrace current best practices for teaching mathematics. To do so, educators must immerse themselves in current research, strive to incorporate the best methods of mathematics instruction, and become educational experts regarding the Common Core State Standards for Elementary School Mathematics. Davis (2015) suggested that in mathematics, a profound understanding of the subject matter taught is a necessary but far from sufficient precondition for providing insightful instruction. Rahman and Lee (2014) explain that the majority of teachers agree that communication in the teaching and learning process is important for enhancing students' understanding. Effective communication, improved students' understanding of the topics being taught, which allowed them to solve high-level questions correctly. According to Lai, Zhu, Chen, and Li (2015), mathematics is among the most objective, logical, and practical academic disciplines. However, in addition to cognitive skills, mathematical problem solving also involves affective factors. Some research supports the conclusion that mathematics anxiety may impede mathematics performance by affecting cognitive process. Mathematics is about problem-solving. Mathematics talks about how to find the best answer of mathematical questions into mathematical statements. The accuracy of answer given by an individual depends on the depth of mathematical knowledge he/she decision (Tella, 2008).

Büyükkarci (2014) stated that evidence indicates that high-quality formative assessment has a powerful impact on student learning. In general terms, formative assessment is aims to help students improve their own learning. In practice, formative assessment is a self-reflective process that aims to promote student achievement.

The formative test is a measurement tool for determining learning quality and encouraging student learning activity. When formative evaluations are conducted frequently during the learning process, student learning outcomes are improved. Badger and Thomas (1992) explain that assessment motivates students to study and includes at least two purposes: encouraging understanding and an increased frequency of studying and determining a method of subject matter learning that the student can understand. Formative assessment (FA) is intended to help learners and teachers track students' progress in an informal way and to take remedial action when learning difficulties emerge (Srivastava, Waghmare, & Vagha, 2015). Formative assessment motivates the student to be more focused and provides an opportunity to monitor different aspects of student learning (Lucas & Spencer, 2014).

Gronlund and Linn (1990) explain that the purposes of assessing learning progress are to (1) establish learning output goals; (2) determine learning needs; (3) monitor learning development and difficulty; (4) assess learning outcomes; and (5) use the evaluation outcome to develop study and learning aids. These purposes mean that the development and implementation of learning aids also serve to influence learning quality and productivity.

According to Arifin (2009), schools generally use only one of many available assessment tools for evaluations: the written test. The written test can take two forms: a subjective test and an objective test. Both tests have different strengths and weaknesses. Essay (subjective) tests require students to organize, interpret, and connect the knowledge that they already have. They also require students to consider, recognize, and improve their existing knowledge while applying a high level of creativity. In contrast, multiple choice (objective) tests examine students' knowledge about specific concepts and skills (Wiersma & Jurs, 1990).

The use of formative assessment tools in the form of essay and multiple choice tests has been predicted to affect students' self-directed learning. Self-directed learning is an intentional psychological activity that students direct and control with the aim of acquiring knowledge and understanding about a specific subject. Students' level of motivation for self-directed learning has an effect on their learning outcomes, which are indicators that the student's knowledge and understanding of a specific subject is improving.

Gronlund and Linn (1990) noted that tests are systematic procedures for measuring behavior or for determining how an individual acts when compared with others or when certain assignments need to be completed. Furthermore, Morrow, Mood, Disch, and Kang (2005) stated that a test is an instrument that is used to measure a specific skill. Such instruments can include written, oral, physiological, and/or mechanical devices. These opinions are supported by Hopkins (1986), who argue that a test is an instrument, tool or procedure that contains assignments that students should address and that provides results that can be used to measure certain aspects of students' knowledge.

Cronbach (1994) and Nitko (2001) stated that a test is a systematic procedure for monitoring and describing one or more student characteristics using a numerical scale or classification scheme and/or a numerical standard or category system. Anastasi and Urbina (1997) wrote that a test is an instrument with objective standardization, and its results can be used broadly (for example, to compare psychological circumstances or individual behaviors). Regarding the function and purpose of tests, Popham (1995) stated that they are useful for diagnosing students' strengths and weaknesses, determining student development, deciding student rankings, and determining the efficacy of further

learning. In fact, Hopkins (1986) cites the final item on this list as the main purpose of testing in terms of quantifying student learning outcomes.

Regarding the function of tests, Gronlund and Linn (1990) stated that they can be classified into four categories: placement tests, formative tests, diagnostic tests and summative tests. A placement test is used to determine the student's ability at the beginning of the learning process; a formative test is used to monitor learning output; a diagnostic test is used to diagnose learning difficulty; and a summative test is used to evaluate achievements.

According to Gronlund and Linn (1990) and Boston (2002), formative assessment is designed to provide feedback to students and teachers during the learning process. This feedback should improve students' awareness of the differences between the purposes of instruction and the knowledge, understanding, and skills that they already have, and it should help them to achieve the purposes of learning. Furthermore, Gregory (2000) noted that formative tests aim to determine the degree to which the student is "formed" after participating in the learning process for a certain period of time. Formative testing is usually conducted throughout the learning program, that is, at the end of a teaching unit.

The subjective test is one of the best test types for determining scores that are affected by other's opinions or assessments. Essay tests require students to formulate their own answers; students cannot select a single correct answer but must answer in their own words. The responses to essay questions should be read one at a time, compared with the correct answer supplied by the test maker, and then scored according to previously provided guidelines. Stiggins (2001) stated that the essay test is a type of written test that consists of questions and requires students to create their own answers.

Hopkins (1986) noted that essay tests can provide insight into several student abilities: (1) the capability to think, analyze, synthesize, and evaluate; (2) the maximum capability for developing thinking skills; (3) the capability to argue/express opinions; (4) the capability for written expression; and (5) the maximum capability for naturally organizing logical thinking. Morrow et al. (2005) observed that the open-question nature of the essay test makes it very effective for measuring students' abilities to organize, analyze, synthesize and assess compared with other tests. Essay tests can also effectively measure participants' arguments and attitudes.

The essay test offers several advantages, such as (1) assessing the mental processes that students use to shape their ideas into correct answers, (2) measuring students' ability to answer a question in their own words, (3) encouraging students to actively learn and to arrange, construct and explain their logical thinking, (4) encouraging students to be courageous when making an argument and to construct their arguments in their own words, and (5) understanding how deeply the student understands and has addressed a problem based on the knowledge taught in the class.

In comparison, multiple choice tests are flexible and of high quality, which allows them to be used for many types of tests and examinations (Gronlund & Linn, 1990). Hopkins (2005) explained that multiple choice tests consist of two parts: incomplete questions or statements and two or more possible answers plus a few other answer options as decoys. The decoys serve to distract the participant when he or she is uncertain of the correct answer.

Cangelosi (2007) said that to facilitate a participant's ability to answer the questions, multiple choice questions should be composed according to certain guidelines, including (a) how the student is expected to select the option, (b) how many options may be selected for each question, (c) whether the answer options include one correct answer or the most

correct answer, and (d) other significant information (for example, whether predicting the answer is allowed).

Dodge (2009) explained that multiple choice tests are used in all areas that require written tests for several reasons: 1) the questions can be analyzed and scored efficiently, quickly and correctly; 2) participants are less hesitant to answer multiple choice tests compared with other test types; 3) questions with two or more possible answers are difficult to answer by guessing; 4) multiple choice tests can be used to measure higher-level skills, such as application and analysis; and 5) they can be used to measure some objective knowledge.

The multiple choice test has several benefits: 1) it can be used to assess several broad learning outcomes; 2) it does not allow the student to write or describe the answer, which minimizes guessing; 3) it prioritizes reading and thinking because the items make minimal demands on memory; 4) it allows students to deduce answers based on those provided; and 5) its use of decoy options can determine whether the student possesses the relevant perception and knowledge (Nitko, 2001).

The factors that affect mathematics learning outcomes in elementary schools are tied to directed learning. Directed learning implies independence. Students in directed learning environments do not depend on someone else for their learning; their efforts are self-initiated, and they must take responsibility in their daily activities. Directed learning is also known as self- or autonomous learning. Seifert and Hoffnung, who were quoted by Eggen and Kauchak (2007), defined autonomy as the ability to control and arrange one's ideas and feelings, to freely act out of self-motivation, and to overcome a lack of self-confidence and self-doubt. Based on the research result, teachers with higher self-efficacy belief are different from those with lower self-efficacy belief. The difference can be seen from the level of efforts, persistence when handling different kinds of students, openness to new ideas and methods, believing in students' achievements and success, and building warm relationship with students rather than with their parents (Nurlu, 2015).

Nash (2014) stated that the idea of self-directed learning originated with Dewey's response to what he saw as the demands of a modern democratic society. His thinking influenced education at the turn of the twentieth century and continues to inform educational policy to this day. For Dewey (2012), the ideal aim of education in a democracy is the creation of self-control that can guide one's personal freedom to choose.

Self-directed learning is the effort to remove oneself from one's parents and find oneself through self-identification, which is the directed development of a stable and independent individual (Good & Brophy, 2007). The directed attitude is not egoistic or selfish; rather, it is a willingness and ability to develop one's own life. Self-directed learning is marked by the individual's ability to adapt his/her behavior, take responsibility, make his/her own decisions, show initiative and creativity, and be able to solve problems without intervention from others. Self-directed learning can occur under a number of conditions, including when: 1) teachers act as facilitators rather than as sources of content; 2) learners are involved in selecting learning resources and strategies, and 3) learners are involved in self-assessment of their learning outcomes (Knowles, 1975).

Petty (2009) described self-directed learning as a humanistic approach in which teachers serve as facilitators. The benefits of independence are that it encourages active learning and extends the responsibility for learning to the student. Furthermore, Scott (2006) and Tavani and Losh (2003) stated that independence/directed learning promotes a child's ability to select correct options, determine his/her options, and be responsible for her/his determination. Independent children have higher self-confidence and motivation, and their behaviors do not depend on anyone else. Van Merriënboer and Sluijsmans

(2009) concluded that instruction that provides a solid base for self-directed learning contains three elements: learners must (a) perform the tasks, (b) assess their task performance, and (c) select future tasks for improving their performance.

Independent learning is recognized as an important predictor of students' academic motivation and achievement. It is a pivotal construct in contemporary accounts of effective academic learning. Independent learning is essential to the learning process through which students direct their acquisition of academic knowledge (Winne, 1995).

Schunk (2005) argued that independent learning is a state or condition in which students pursue learning activities on their own without relying on others; this condition or state is always consistent, and self-directed learners are eager to learn anywhere. Independent learning is mostly driven by the student's own initiative, choice and self-learning responsibility.

Zimmerman (1990) defined independent learning as a process of learning that results from the influence of thoughts, feelings, strategy, and self-behavior that are oriented toward achieving a goal. Thomas (1993) suggested that independent learning is an unrestricted and responsible student behavior geared toward determining learning objectives, planning and implementing, and maintaining and assessing the results of learning activities without relying on others. Zimmerman (2008) viewed independent learning as a proactive process rather than a reactive event that happens to students due to impersonal forces such as teaching. Based on the above opinions, independent learning depends on the student's attitude and ability to complete learning activities independently or alone and be responsible for achieving a goal. Independent learning in mathematics results from a student's attitude and ability to learn mathematics independently, master a competency and be responsible for completing the task with reduced guidance from others (Cleary & Chen, 2009).

During the self-learning process, students require independence if the learning process able to progress properly and maximize student achievement. Metallidou and Vlachou (2010) said that self-learning is learning in which students are allowed to determine the learning goals; plan the learning process and the strategies using learning resources that he or she has chosen; make academic decisions; and perform activities that contribute to the achievement of learning goals.

Banarjee and Kumar (2014) stated that independent learning has an effect on academic achievement, including in mathematics. This effect occurs because children begin to believe that they have the ability, discipline, and enthusiasm to pursue achievement; as a result, they do not feel inferior and are ready to solve any problems that arise, and this attitude increases achievement. According to Banarjee and Kumar (2014), independent learning and academic achievement are positively correlated, and there is a significant positive relationship between independent learning and academic achievement in science. In the context of mathematics classrooms, students are challenged to solve tasks, to conceptualize their own opinions and to adapt strategies to task demands (Kramarski & Revach, 2009).

The research by Lopez Vargas, Hederich-Martinez, and Camargo Uribe (2012), Ocak and Yamaç (2013), Camahalan (2006), Metallidou and Vlachou (2010) and Çiftçi and Koza (2015) indicates that to improve and maintain mathematics learning outcomes, it is necessary to consider factors related to the students' independence with mathematics. When a student's level of independence is high, it will affect the results of that student's mathematics learning. The greater the student's independence, the higher the mathematics learning outcomes; conversely, the lower the student's independence, the lower his or her mathematics learning outcomes will be.

The research conducted by Yang and Li (2013) assessing the use of animated self-directed learning activity modules for children's number sense development found that even without a teacher's instruction in class, children can develop number sense through self-directed learning.

Research conducted by Munasco (2013) in Banda Aceh concluded that students' physics learning outcomes are better when essay tests are used than when multiple choice tests are used, after controlling for initial capability. The formative testing research conducted by Putri and Indra (2010) with elementary students in Palembang concluded that the students who were given essay tests showed higher mathematics learning outcomes than the students who were given multiple choice tests, after controlling for initial capability.

Method

This research aimed to determine the effects of and interactions between formative testing and self-directed learning and their relationship with the mathematics learning outcomes of elementary students.

The research was conducted in a public elementary school in central Jakarta. The research subjects were students who were in their first semester of fourth grade. The research began as an experiment to examine the validity of an instrument (which took the form of essay tests and multiple choice tests) for assessing learning outcomes.

The research method used was an experimental method with a 2 x 2 treatment-by-level design. The variables in this research were dependent variables (the formative tests), an independent variable (mathematics learning outcomes), and an attribute variable (self-directed learning). The formative tests (A) included two forms: a formative essay test (A₁) and a multiple choice test (A₂). Student self-directed learning (B) was classified as either high (B₁) or low (B₂). There were four groups tested: a group of essay tests and students with high self-directed learning (A₁B₁), a group of multiple choice tests and students with high self-directed learning (A₂B₁), a group of essay tests and students with low self-directed learning (A₁B₂), and a group of multiple choice tests and students with low self-directed learning (A₂B₂).

Table 2. *Design Treatment by Level*

Self-directed learning (B)	Formative test (A)	
	Essay (A₁)	Multiple choice (A₂)
High (B ₁)	A ₁ B ₁	A ₂ B ₁
Low (B ₂)	A ₁ B ₂	A ₂ B ₂

The two-treatment method was designed to determine the effect of the treatments on mathematics learning outcomes. The examined treatments were formative tests and independent/self-directed learning. The formative test treatment involved giving the students formative essay tests and formative multiple choice tests. Formative essay tests are tests that require students to create their own answers. Multiple choice tests provide several possible answers and require the student to determine which is correct. The formative multiple choice test that was used in this study provided four answer options for each question. The students' self-direction levels were assessed via a questionnaire.

The research sample pool comprised all of the students in the fourth grade at a public elementary school in central Jakarta. The sample was selected using a random sampling technique. Two classes were used as sampling sources for this research. Two classes were selected so that one class could take the formative essay test and the other could take the formative multiple choice test. Grade IV A was assigned to take the formative essay test,

and Grade IV B was assigned to take the formative multiple choice test. The students in Grade IV A and Grade IV B had the same relative academic capabilities because neither group of students was categorized as having superior skills. The sample population comprised 72 students.

To measure self-directed learning, the students were grouped according to whether they reported having high or low levels of directed learning. The sample included 40 students, 20 of whom took essay tests and 20 of whom took multiple choice tests. Ten students in each group were chosen randomly for the self-directed learning assessment.

The experiment was conducted using formative essay tests and multiple choice tests. The students who took the formative essay tests were tested 3 consecutive times at the end of each lesson, as were the students who took the formative multiple choice tests.

Furthermore, to avoid differences in test scores that may have been caused by the formative tests that were used for practice, the final test for each student group included a bundle of essay or multiple choice tests, as appropriate for the group. The data collection instruments were (1) a mathematics learning outcome instrument and (2) a student-directed learning instrument. The student-directed learning data were measured using a non-test instrument that used frequency scales with five options. These data were collected before the treatment. The data were intended to group the respondents according to their self-directed learning levels (high and low). The mathematics learning outcome data were measured using a learning outcomes test, which consisted of essay and multiple choice tests. Both tests were developed by the researcher and tested on students with characteristics similar to those of the research respondents.

Validity testing was conducted using tests of content validity and construct validity. Construct validity was tested using expert judgment. Content validity was tested by referring to the elementary school curriculum for fourth grade. The multiple choice questionnaire used a dichotomy formula based on biserial points. The tests were created using Microsoft Excel 2013 software (Microsoft Operations Pte, Ltd., Singapore, Reagents: Sistech Kharisma Company). The validity of each questionnaire was determined by comparing the correlation coefficient (r_{value}) with the critical biserial correlation number (r_{table}) based on a significance level of 5%, as follows: 1) if $r_{\text{item}} > r_{\text{table}}$ and $\alpha = 0.05$, then the item was considered valid; 2) if $r_{\text{item}} \leq r_{\text{table}}$ and $\alpha = 0.05$, then the item was considered invalid. Based on this calculation, 16 questions on the multiple choice test had a value of $r_{\text{phi}} > 0.355$ and a significance level of $\alpha = 0.05$. The validity test for the essay tests, based on the product moment formula, showed that the entire test (7 questions) had a value of $r_{\text{value}} > r_{\text{table}}$ (0.355) and a significance level of $\alpha = 0.05$. To determine the reliability of instrument, the Hoyt formula was used.

Before the hypothesis was tested, we conducted an analysis requirement test, which consisted of a normality test and a homogeneity test. The normality test used the Lilliefors formula. The data were found to have a normal distribution when $L_0 < L_{\text{table}}$ with a significance level of 0.05. The variance homogeneity test of Group Y for X_1 and X_2 was intended to test the similarity of two population variants with normal distribution patterns. The homogeneity of the data were tested using Bartlett's test. The data were homogeneous if $\chi^2_{\text{count}} < \chi^2_{\text{table}}$, with a significance level of 0.05.

Results

The hypothesis was tested using a two-line variance analysis to examine the interactions, and then a one-line analysis was used to test the hypothesis regarding the simple effect. The two-line analysis results were used to test the hypothesis regarding the main effect, which was that the average learning outcomes of the student groups that took the formative essay tests would differ from those of the student groups that took the formative multiple choice tests. Furthermore, the hypothesis about the effect of the interaction between the formative test treatments and self-directed learning on mathematics learning outcomes was tested. The results for these hypotheses are summarized in Table 3 below.

Table 3. Two-Line ANOVA Table for Mathematics Learning Results with $\alpha = 0.05$ and $\alpha = 0.01$

Variance source	JK	Db	RJK	F value	F table	
					$\alpha = 0.05$	$\alpha = 0.01$
Inter A	114.568	1	114.568	5.38*	4.08	7.31
Inter B	589.114	1	589.114	27.66**	4.08	7.31
Interaction A x B	134.750	1	134.750	6.32*	4.08	7.31
Inside/within	852.000	40	21.300			
Total	1690.432	43				

The results of the two-line ANOVA above can be explained as follows:

1) *The mathematics learning results of the students who were given the formative essay test treatment differed from those of the students who were given the formative multiple choice test treatment.*

Based on the ANOVA results in Table 3 above, $F_{\text{value}} = 5.38$, whereas F_{table} with $\alpha = 0.05$ is 4.08. Thus, $F_{\text{value}} > F_{\text{table}}$, which indicates that H_0 must be rejected because H_1 is accepted. Thus, there is a difference in the average mathematics learning outcomes of the students who received the formative essay test treatment and those who received the formative multiple choice test treatment. The average mathematics learning results for the students who took the formative tests was 71.60, whereas the students who took the formative multiple choice had an average mathematics learning result of 66.00.

2) *The mathematics learning outcomes differed between the students with high self-directed learning and those with low self-directed learning.*

The ANOVA results in Table 3 above show that $F_{\text{value}} = 27.66$, whereas the F_{table} at $\alpha = 0.01$ is 7.31. Thus, $F_{\text{value}} > F_{\text{table}}$, which shows that H_0 is rejected because H_1 is accepted. Therefore, there is a difference in the mathematics learning outcome averages between the group of students with high self-directed learning and the group with low self-directed learning. According to the mathematics learning test results, students with a high level of independence had an average formative essay test score of 83.4, while the average score on the formative multiple choice test was 70.7.

3) *The formative test type and self-directed learning have an interaction effect on mathematics learning outcomes.*

The ANOVA results above show that the F_{value} had an interaction effect $AB = 6.32$, whereas F_{table} at $\alpha = .05$ was 4.08. Thus, $F_{\text{value}} > F_{\text{table}}$, indicating that H_0 was rejected and H_1 was accepted. It may be concluded that there is an interaction effect between the formative test type and self-directed learning. The effect of the interaction between the type of formative test and self-directed learning on mathematics learning outcomes can be seen in Figure 1.

The average F between the four groups of cells was 39.78, while the F_{table} at the level of $\alpha=0.05$ is 2.84. Thus, $F_{count} > F_{table}$, meaning that H_0 is rejected and H_1 is accepted, indicating that on average, there were differences among the four groups of cells.

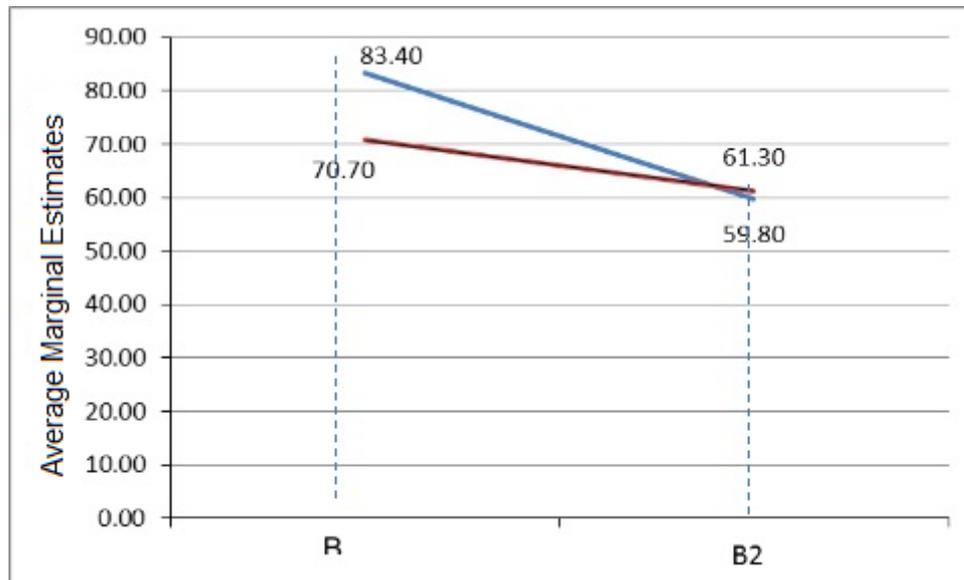


Figure 1. The interaction between formative test type and self-directed learning.

The intersection of the two lines shows that there is an interaction between the two variables toward the independent variable. In other words, the interaction means that the effect of the formative test type on the mathematics learning outcomes depends on the student's level of self-directed learning.

The intersection of the two lines also shows that there is an interaction between the two independent variables. In other words, the interaction means that the effect of the formative test type on mathematics learning outcomes depends on the student's level of self-directed learning.

As a consequence of the interaction, the two-line ANOVA must be continued with a Dunnett's t -test. The test was performed to examine the simple effect, that is, to examine the difference in mathematics learning outcomes between the students with high self-direction who received the formative essay test and those who received the formative multiple choice test and the difference in mathematics learning outcomes between the students with low self-direction who received the formative essay tests and those who received the formative multiple choice tests.

4) *The mathematics learning outcomes differed between the students with high levels of self-directed learning who received the formative essay tests and those who received the formative multiple choice tests.*

The results of Dunnett's t -test showed that the average mathematics learning outcome scores differed for the students with high levels of self-directed learning who received the formative essays and those who received the formative multiple choice tests. The t_{value} was 3.42, whereas the value for t_{table} at level $\alpha = 0.01$ was 2.43; thus, $t_{value} > t_{table}$. As a result, H_0 is accepted, and H_1 is rejected, indicating that the mathematics learning outcomes differed between the highly self-directed students who received the formative essay tests and those who received the formative multiple choice tests.

The data obtained for the mathematics learning outcomes of the students with high levels of self-directed learning showed that the average scores for those who received the

formative essay tests were approximately 41.9, whereas the average scores of those who received the formative multiple choice tests were approximately 35.18, indicating that the average scores of the students who were given formative essay tests were higher than those of the students who were given formative multiple choice tests. Consequently, the learning output of the students with high levels of self-directed learning who were given formative essay tests was higher than that of the highly self-directed students who were given formative multiple choice tests. Therefore, it may be said that giving formative essay tests to these students could increase their learning results.

5) The mathematics learning results differed between the students with low levels of self-directed learning who received the formative essay tests and those who received the formative multiple choice tests.

The Dunnett's t-test results, which aimed to differentiate the average mathematics learning results of the students with low levels of self-directed learning who were given formative essay tests from the learning results of those who were given the formative multiple choice tests, showed that the t_{value} was roughly - 0.138, whereas the t_{table} score at $\alpha=0.05$ is -1.68. Therefore, $t_{\text{value}} > t_{\text{table}}$, and H_0 was accepted or H_1 was rejected. This result indicates that there was a difference in the mathematics learning outcomes of the students with low levels of self-directed learning who were given the formative essay tests and those who were given the formative multiple choice test.

The mathematics learning outcome data of the students with low self-directed learning showed that the average score of the students who received the formative essay tests was approximately 31.09, whereas the average score of those who received the formative multiple choice tests was 31.36. These findings indicate that the learning results of those who were given the formative essay tests were lower than the results of those who were given the formative multiple choice tests. The findings also show that the learning result for students with low levels of self-directed learning were lower when they were given formative essay tests than when they were given the formative multiple choice tests. Thus, for students with low levels of self-directed learning, formative multiple choice tests could improve their learning results.

Discussion

The results of our research and statistical analyses can be explained by noting that the variation in the mathematics learning outcomes of the students who were given the formative essay tests and those who were given the formative multiple choice tests. The mathematics learning outcome results showed that the students who took the formative essay tests scored higher than the students who took the formative multiple choice tests. This finding indicates that the type of formative test has an overall effect, regardless of whether the students have high or low levels of self-directed learning. This evidence is strengthened by similar research conducted by Munasco (2013), in which students who were given essay tests scored higher than students who were given multiple choice tests.

The research hypothesis, which stated that mathematics learning outcomes would be higher for students who were given formative essay tests than for those who were given formative multiple choice tests, was empirically examined via data analysis. The findings are explained by comparing the characteristics of both types of formative tests. As the measured learning outcomes show, the formative essay test is more effective for measuring high levels of cognitive ability, but it lacks the ability to effectively measure low levels of cognitive ability. In contrast, the formative multiple choice test is more effective for measuring low and moderate levels of cognitive ability but lacks effectiveness for measuring high levels of cognitive ability. Regarding the scope of the material, the essay test has a small scope and a small number of questions compared to the multiple choice

test. The formative multiple choice test covers a large scope of material and can represent an entire subject. However, as the factors that affected the test results show, students' scores on formative essay tests are affected by their ability to analyze, evaluate and argue their ideas on paper.

Theoretically, a formative test only needs to utilize one type of answer structure to allow students to describe, explain, compare, and reason by answering questions in their own words and by using their analytical thinking and interpretation skills to address the questions. The essay test gives students the freedom to express their ideas, which could encourage them to express an argument/opinion through their analytical writing and to express their answer directly in a way that could increase the quality of the answer and the score. This means that all students who are given essay tests can use their imagination and explore their ideas to the best of their ability. Through formative essay tests, students are motivated to develop their higher-order thinking and express it creatively.

Formative multiple choice tests provide different advantages. They are good for measuring students' abilities because they do not allow students to use their imagination. Multiple choice test results are affected by the student's ability to understand the concept, which can be viewed as a learning result. A formative multiple choice test can encourage students to remember, interpret and recall other people's ideas. Consequently, it could be said that the students who were given formative essay tests showed higher mathematics learning results than the students who were given multiple choice formative tests.

These facts support the opinion that the average mathematics learning outcome scores on formative essay tests are higher than those on formative multiple choice tests; consequently, these findings are evidence that the formative essay test treatment is more effective than the formative multiple choice test treatment for increasing student mathematics learning results. This finding is in line with the previously discussed characteristics of each formative test type. Consequently, the stated hypothesis is scientifically proven: formative essay tests improve student learning results more than formative multiple choice tests.

The second research hypothesis stated that there is an interaction effect between formative essay tests, formative multiple choice tests, and student self-directed learning on mathematics learning results, which shows the connectivity between the variables. The empirically tested data for the formative essay and multiple choice tests showed that the tests have different levels of effectiveness for measuring student learning results, depending on the student's level of self-directed learning. The data are strengthened by the research by Gavriel (2013), which stated that the effectiveness of formative assessment depends on the level of student self-directed learning. According to Oyediji and Okwilagwe (2015), self-directed learning is a process through which individuals consciously take responsibility and initiative, with or without the help of others, to determine their learning needs, formulate learning goals, identify resources for learning, select and implement learning strategies and evaluate learning outcomes within a given framework, thereby becoming their own learning agents (Costa & Kallick, 2004).

The average mathematics learning results, according to the level of student self-directed learning, were higher for the students with high levels of self-directedness who were given formative essay tests than for those who were given formative multiple choice tests. This finding shows that students who have high levels of self-directed learning can respond more effectively to formative essay tests than to multiple choice tests and that the use of essay tests could increase their learning results. However, the average score of students who had low levels of self-directed learning and received formative essay tests was lower than that of the students who were given multiple choice tests. This finding

indicates that students with low levels of self-directed learning work better with multiple choice tests and that the use of such tests could improve their learning results.

The third hypothesis is that the results for students with high levels of self-directed learning who were given formative essay tests would differ from the results for those with high levels of self-directed learning who were given formative multiple choice tests. The analytical results showed a difference in learning results between the students who were given formative essay tests and the students who were given formative multiple choice tests among the groups with high levels of self-directed learning. That is, the students with high levels of self-directed learning perform better on formative essay tests than on formative multiple choice tests. This result agrees with the opinion stated by Hapsari (2013), who said that students with high performance motivation show higher mathematics learning results on formative essay tests than on formative multiple choice tests. If self-directed learning is connected to formative essay tests, which are used for formative evaluation, then the characteristics of high self-directed learning are appropriately matched with the characteristics of essay tests; that is, the opportunity to apply individualized problem-solving skills, implement formulas and analyses, and take responsibility for one's assignments. Those characteristics indicate that the formative essay test requires critical thinking to find the answer. Thus, the formative essay test's characteristics are appropriately matched with a high level of self-directed learning, which reflects an ability to complete activities without depending on other people and having good self-confidence and initiative.

The fourth hypothesis stated that students with low levels of self-directed learning who were given formative essay tests would have lower mathematics learning outcomes compared with the students with low levels of self-directed learning who were given multiple choice tests. The data showed that the students with low levels of self-directed learning had lower scores on the formative essay tests. This occurred because these students did not have the capability to solve the problem, the self-confidence to answer it correctly, or the analytical thinking skills to arrive at an answer. This type of student tends to lack initiative in the learning process, is more likely to have a fixed approach to the subject, depends on teachers' explanations, is apprehensive about unclear ideas, and lacks the encouragement or effort to find the correct answer. Students who have low self-directed learning find it more difficult to solve problems on formative essay tests than on formative multiple choice tests. This is because students with low self-directed learning find it difficult to express their ideas, implement formulas and make decisions, which are required characteristics of the formative essay test. Embo, Driessen, Valcke and Van der Vleuten (2010) suggest that the integration of feedback and assessment through a clearly defined learning and assessment instrument is a potentially valuable method for promoting self-directed learning and formative assessment.

Conclusions

Based on the data analysis and statistical calculations, this research draws the following conclusions. First, students who were given formative essay tests and those who were given formative multiple choice tests had different mathematics learning outcomes. The students who received formative essay tests scored higher than those who received formative multiple choice tests. Second, there was an interaction effect between formative tests and self-directed learning on mathematics learning outcomes, which means that the effect of the interaction between formative tests and self-directed learning on student mathematics results in elementary school depends on the level (high or low) of self-directed learning. Third, students who had high levels of self-directed learning and were given formative essay tests had higher mathematics learning outcomes compared with those who were given formative multiple choice tests. Finally, among the students who had low levels of self-

directed learning, there was no difference in mathematics learning results between those who were given formative essay tests and those who were given formative multiple choice tests.

Implications

Based on the research findings, the direct implications for mathematics learning performance are as follows:

- 1) Our research findings indicate that the type of formative test used has an effect on the learning results. Thus, to improve students' learning results, teachers should use different types of formative tests in the classroom, and students should be able to complete the assignments and obtain optimal scores.
- 2) The research findings indicate that students' self-directed learning has an effect on the learning results. This finding implies that during the learning process, a student with a high degree of self-direction will actively participate in every learning activity in the classroom, with optimal learning results.
- 3) Given the appropriate materials and methods, allowing self-direction in the classroom will develop students who are active, creative, and critical and will increase students' learning motivation, which will significantly affect their mathematics learning results.



Acknowledgements. The authors would like to thank Jody C; Shannon T; Jodi H; and Stacie M for their helpful comments and useful suggestions regarding an earlier draft of this manuscript.

References

- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall Inc.
- Arifin, Z. (2009). *Evaluasi pembelajaran* [Learning evaluation]. Bandung, Indonesia: PT Remaja Rosdakarya (ROSDA).
- Badger, E., & Thomas, B. (1992). Open-ended questions in reading. Practical assessment. *Research & Evaluation, 3*(4), 1.
- Banarjee, P., & Kumar, K. (2014). A study on self regulated learning and academic achievement among the science graduate student. *International Journal of Multidisciplinary Approach & Studies, 1*(6), 329-342.
- Boston, C. (2002). The concept of formative assessment. Practical assessment. *Research & Evaluation, 8*(9), 1.
- Büyükkarci, K. (2014). Assessment beliefs and practices of language teachers in primary education. *International Journal of Instruction, 7*(1), 108.
- Camahalan, F. M. G. (2006). Effects of self regulated learning on mathematics achievement of selected southeast Asian children. *Journal of Instructional Psychology, 33*(3), 194.
- Cangelosi, J. S. (2007). *Classroom management strategies: Gaining and maintaining students' cooperation* (6th ed.). Hoboken, NJ: Wiley.

- Çiftçi, Ş., & Koza. (2015). Effects of secondary school students' perceptions of mathematics education quality on mathematics anxiety and achievement. *Educational Sciences: Theory & Practice*, 15(6), 1487.
- Cleary, T. J., & Chen, P. P. (2009). Self-regulation, motivation, and math achievement in middle school: Variations across grade level and math context. *Journal of School Psychology*, 47(5), 291-314. doi:10.1016/j.jsp.2009.04.002
- Costa, A. L., & Kallick, B. (2004). Launching self-directed learners. *Educational Leadership*, 62(1), 51-52.
- Cronbach, L. J. (1994). *Essentials of psychological testing*. New York, NY: Harper & Row.
- Davis, B. (2015). The mathematics that secondary teachers (need to) know. *Revista Española de Pedagogía*, 73(261), 321-342.
- Dewey, J. (2012). *Democracy and education*. LaVergne, TN: Simon & Brown.
- Dodge, J. (2009). 25 quick formative assessments for a differentiated classroom. Sydney, Australia: Scholastic Teaching Resources.
- Eggen, P., & Kauchak, D. (2007). *Educational psychology: Window on classrooms*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Embo, M. P., Driessen, E. W., Valcke, M. & Van der Vleuten, C. P. (2010). Assessment and feedback to facilitate self-directed learning in clinical practice of Midwifery students. *Medical Teacher*, 32(7), e263-e269.
- Gavriel, J. (2013). Assessment for learning: A wider (classroom-researched) perspective is important for formative assessment and self-directed learning in general practice. *Education for Primary Care*, 24(2), 93-96.
- Good, T. L., & Brophy, J. E. (2007). *Looking in classrooms* (10th ed.). Needham Heights, MA: Allyn & Bacon.
- Gregory, R. J. (2000). *Psychological testing: History, principles and applications* (3rd ed.). Boston, MA: Allyn and Bacon.
- Gronlund, N. E., & Linn, D. R. L. (1990). *Measurement and evaluation in teaching* (6th ed.). New York, NY: Macmillan Publishing Company.
- Hapsari, S. A. (2013). Kontribusi kemandirian terhadap hasil belajar matematika ditinjau dari fasilitas belajar dan jarak tempat tinggal siswa [The contribution of independence of the results of learn math in terms of facilities learning and the distance between their houses students]. Retrieved from http://eprints.ums.ac.id/23458/12/NASKAH_PUBLIKASI.pdf
- Hopkins, David (1986). *Evaluation For School Development*. United Kingdom: Open University Press Publishers.
- Hopkins, K. (2005). *Educational and psychological measurement and evaluation* (7th ed.). Englewood Cliffs, NJ: Prentice Hall.
- Knowles, M. S. (1975). *Self-directed learning: A guide for learners and teachers*. New York, NY: Association Press.
- Kramarski, B., & Revach, T. (2009). The challenge of self-regulated learning in mathematics teachers' professional training. *Educational Studies in Mathematics*, 72(3), 379-399. doi:10.1007/s10649-009-9204-2
- López Vargas, O., Hederich-Martinez, C., & Camargo Uribe, Á. (2012). Mathematics achievement, self regulated learning and cognitive style. *Suma Psicológica*, 19(2), 39-50.
- Lai, Y., Zhu, X., Chen, Y., & Li, Y. (2015). Effects of mathematics anxiety and mathematical metacognition on word problem solving in children with and without mathematical learning difficulties. *PLoS One*, 10(6), e0130570. doi:10.1371/journal.pone.0130570

- Metallidou, P., & Vlachou, A. (2010). Children's self regulated learning profile in language and mathematics: The role of task value beliefs. *Psychology in the Schools, 47*(8), 780-788.
- Morrow Jr., J. R., Mood, D., Disch, J., & Kang, M. (2005). *Measurement and evaluation in human performance*. Champaign, IL: Human Kinetics Publishing House.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2011). International student achievement in the TIMSS mathematics content and cognitive domains. In TIMSS 2011 international results in mathematics (pp. 139-171). Retrieved from http://timssandpirls.bc.edu/timss2011/downloads/T11_IR_M_Chapter3.pdf
- Munasco, M. (2013). The formative test form, teachers' qualification, and students' achievement in physics by controlling the students' prior knowledge. *Jurnal Evaluasi Pendidikan, 4*(1), 37-51. Retrieved from <http://www.library.gunadarma.ac.id/journal/view/11011>
- Nash, C. (2014). Founders' continuing roles in schools supporting self-directed learning. *Interchange, 45*(1/2), 43-57.
- Nitko, A. J. (2001). *Educational assessment of student* (3rd ed.). Upper Saddle River, NJ: Merrill.
- Nurlu, Ö. (2015). Investigation of Teachers' Mathematics Teaching Self-efficacy. *International Electronic Journal of Elementary Education, 8*(1), 24.
- Ocak, G., & Yamaç, A. (2013). Examination of the relationships between fifth graders' self-regulated learning strategies, motivational beliefs, attitudes, and achievement. *Educational Sciences: Theory & Practice, 13*(1), 380-387.
- Osborne, J. H. (2015). The charge is ours: Teaching elementary school mathematics through best practice and common core initiatives. *Delta Kappa Gamma Bulletin, 82*(1), 23-25.
- Oyediji, O., & Okwilagwe, E. (2015). Investigating the effects of self-directed learning and collaborative methods on junior secondary school students social studies learning outcomes in Oyo state, Nigeria. *Problems of Education in the 21st Century, 64*, 38-52.
- Petty, G. (2009). *Teaching today: A practical guide* (4th ed.). Cheltenham, England: Nelson Thornes.
- Popham, W. J. (1995). *Classroom assessment: What teachers need to know*. Needham Heights, MA: Allyn and Bacon.
- Putri, R., & Indra, I. (2010). Pengaruh pendekatan pembelajaran dan bentuk tes formatif terhadap hasil belajar matematika dengan mengontrol intelegensi siswa sd di Palembang [The effect of learning approach dan formative test to mathematic learning outcome: Study in Palembang public elementary student] (Postgraduate dissertation, Jakarta State University, Indonesia).
- Rahman, N. A. A. A., & Lee, M. F. N. (2014). Communication in teaching and learning mathematics: Teachers' perspective. *AIP Conference Proceedings, 1605*, 730-733. doi:10.1063/1.4887680
- Schunk, D. H. (2005). Self-regulated learning: The educational legacy of Paul R. Pintrich. *Educational Psychologist, 40*(2), 85-94. doi:10.1207/s15326985ep4002_3.
- Scott, K. W. (2006). Self-directed learners' concept of self as learner: Congruous autonomy. *International Journal of Self-Directed Learning, 3*(2), 1-13.
- Srivastava, T., Waghmare, L. S., & Vagha, S. J. (2015). Revisiting feedback practices in formative assessment of Indian Medical Schools. *National Journal of Physiology, Pharmacy & Pharmacology, 5*(1), 1-4. doi:10.5455/njppp.2014.4.270620141
- Stiggins, R. (2001). *Student-involved classroom assessment* (3rd. ed). Upper Saddle River, NJ: Merrill Prentice-Hall.
- Tavani, C. M., & Losh, S. C. (2003). Motivation, self confident and expectations as predictors of the academic performances among our high school students. *Child Study Journal, 33*(3), 141-152.
- Thomas, J. W. (1993). Promoting independent learning in the middle grades: The role of instructional support practices. *Elementary School Journal, 93*(5) 575-591. doi:10.1086/461741

- Tella, A. (2008). Teacher Variables As Predictors of Academic Achievement of Primary School Pupils Mathematics, *International Electronic Journal of Elementary Education*, 1(1) 17.
- Van Merriënboer, J. J. G., & Sluijsmans, D. M. A. (2009). Toward a synthesis of cognitive load theory, four-component instructional design, and self-directed learning. *Educational Psychology Review*, 21(1), 55–66. doi:10.1007/s10648-008-9092-5
- Wiersma, W., & Jurs, S. G. (1990). *Educational measurement and testing* (2nd ed.). Retrieved from <http://www.clr.ui.ac.id/wp-content/uploads/2010/10/Educational-Measurement-and-Testing.pdf>
- Winne, P. H. (1995). Inherent details in self-regulated learning. *Journal of Educational Psychology*, 30(4), 173–187. doi:10.1207/s15326985ep3004_2
- Yang, D.-C., & Li, M. N. (2013). Assessment of animated self-directed learning activities modules for children's number sense development. *Educational Technology & Society*, 16(3), 44-58.
- Zimmerman, B. J. (1990). A study on self-regulated learning and academic achievement: An overview. *Journal of Educational Psychologist*, 25(1), 13.
- Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, 45(1), 166-183. doi:10.3102/0002831207312909

www.iejee.com

This page is intentionally left blank