Science Education International
33(4), 383-391
https://doi.org/10.33828/sei.v33.i4.5

International Council of Association for Science Education

ORIGINAL ARTICLE

Effects of Technology-integrated Formative Assessment on Students' Retention of Conceptual and Procedural Knowledge in Chemical Equilibrium Concepts

Tadesse Hagos[1]*, Dereje Andargie[2]

[1] Department of Science and Mathematics Education, Addis Ababa University, Addis Ababa, Ethiopia, [2]Department of Chemistry, Instituted of Education, Debre Berhan University, Debre Berhan, Ethiopia

***Corresponding Author:** tadessehagos36@gmail.com

ABSTRACT

The main objective of this study was to evaluate the impact of technology-integrated formative assessment strategies on students' retention of conceptual and procedural knowledge in chemical equilibrium concepts. A quasi-experimental control group interrupted time series design was employed. Data were collected from 132 students, who were selected using a random sampling technique from three governmental secondary schools. Two experimental groups and one comparison group were involved in the study. A series of chemical equilibrium conceptual and procedural tests were used to collect data. A mixed model analysis of variance (ANOVA) was used to analyze the test scores. The three groups had varied mean gain scores in conceptual and procedural examination scores as repeated measurements of time, as can be observed from the descriptive statistics. The two-way mixed ANOVA findings revealed that there was a significant main effect between groups on conceptual test scores. Furthermore, on conceptual knowledge test scores, there was a significant interaction between time point and intervention groups. On procedural examination scores, there was also a significant difference between the 3-time points and intervention groups. A significant interaction between time points and groups was also observed. Given that there was a significant difference between the pre-test, post-test, and follow-up post-test scores in the experimental group, with the difference favoring the post-test scores, it can be concluded that the technology-integrated formative assessment strategy is successful in raising secondary school students' conceptual and procedural test scores.

KEY WORDS: Conceptual knowledge; formative assessment; knowledge; procedural; technology

INTRODUCTION

The 21st century has been described as a period of transformation and reform (Barak, 2017). Many aspects of society, including education, have been influenced by this new way of sharing and distributing knowledge. Learners are using technology to find answers to problems, explore new areas, and discuss ideas with others, which has changed the way they learn (Jahnke et al., 2020). The use of educational technology in the classroom has the potential to change how students and teachers learn and teach (Manuguerra and Petocz, 2011). Hickey et al. (2012) showed how technology may help students learn more about science, engage in a range of scientific activities, and stay involved in those practices. The assessment process has also been changed by technology. It is commonly used for formative assessments and a variety of other purposes, such as reaching more students (Feldman and Capobianco, 2008), motivating and engaging students (Kay and Knaack, 2009; Tan and Towndrow, 2009), modifying lessons (Gerard et al., 2016), providing feedback and scaffolding (Furtak et al., 2016).

Teachers may be able to address students' misconceptions using technology integrated formative assessment (Bhagat and Spector, 2017). Individual students can now receive real-time, personalized feedback suited to their performance thanks to technology improvements (Lu and Cutumisu, 2021). Schools are increasingly integrating such technology into the classroom to provide real-time feedback to students (Tissenbaum and Slotta, 2019). A range of digital-based devices that enhance students' learning have also been created in recent years. Assessment tools are generally included in the study, which provide teachers and students with rapid feedback.

Technology, according to Quellmalz et al. (2013), can help teachers gather resources by providing online databases that link curriculum and national standards; it can provide teachers with a range of assessment tasks and items to embed within lessons and units; and it can help teachers assess complex and dynamic aspects of cognition and performance through technological tools. According to Looney (2019), several of the new technologies incorporate various methods for assessing student performance, including rapid assessment of student understanding; timely and targeted feedback; interactive learning; assessment of higher-order skills; and tracking of students' learning in various contexts and over time.

Science Education International
33(4), 383-391
https://doi.org/10.33828/sei.v33.i4.5

Assessment with technology integration has the ability to provide formative assessment and its advantages to more teachers, students, and classrooms in a timely and useful manner (Elmahdi et al., 2018; Spector et al., 2016). By enabling more rapid feedback, showing feedback in easily usable ways, and giving new opportunities for assessing student understanding of scientific phenomena in dynamic and interactive ways, technology can help educators in effectively implementing formative assessment (Gobert et al., 2013). Even on open-ended, higher-order thinking skills tasks, technology-based systems that monitor students' actions in a non-intrusive fashion can react in real-time based on formative data to scaffold student development (Adesina, 2017). Many scholars believe that technology may help teachers in gathering and evaluating formative data, which can be tough for them (Hsu and Lin, 2020). However, there is not enough evidence on how technology-enhanced formative assessment might help students learn more effectively (Bhagat and Spector, 2017).

Statement of the Problem

Chemistry is an essential part of science because it allows students to understand what is going on around them (Wells, 2002). While abstract concepts are widespread in chemistry, they are essential for the future learning in chemistry and other disciplines (Taber, 2002). These abstract concepts are significant since further chemical concepts or theories will be difficult to understand if the learner does not understanding these basic concepts (Gabel, 1999; Nakiboglu, 2008; Sirhan, 2007). One of the reasons students struggle to understand chemical concepts is a lack of conceptual and procedural knowledge (Habiddin and Page, 2021). To properly respond to any challenge, students must use both conceptual and procedural knowledge (Chen et al., 2021).

Students' poor understanding of both conceptual and procedural problems in chemistry has been blamed on alternative conceptions (Quílez, 2004). In recent years, several types of research have been undertaken on how to increase student learning and knowledge of chemistry. Secondary school students demand student-centered learning settings that encourage and inspire them to improve and create a wide range of conceptual, procedural, and metacognitive knowledge, as well as a wider range of cognitive processes (Anderson and Krathwohl, 2001).

Educators believe that formative assessment strategies are effective in improving students' conceptual and procedural knowledge because teachers can help identify students' misconceptions, make these misconceptions visible to the learners, and implement instructional strategies based on the feedback they receive from the students to address their learning needs (Bernal-Ballen and Ladino-Ospina, 2019; Wei, 2011). Furthermore, it is commonly understood among the educational community that having effective and efficient formative assessments is extremely important, making them a critical component of learning (Paiva et al., 2020; Robertson et al., 2019). As indicated by cognitive science research that emphasizes integrated conceptual and procedural knowledge, technology has a critical role in learning by facilitating the implementation of formative assessment strategies.

However, efforts to implement formative assessment in many countries, including Ethiopia, are hindered by several challenges that contribute to ineffective practices. The time-consuming nature of formative assessment strategies and the time restrictions of class sessions were some of the problems that led to poor practices, making it difficult for teachers to incorporate these strategies into their instruction. Many studies recommend incorporating technology into formative assessment design to overcome challenges such as time constraints, large classrooms with diverse students, and a broad curriculum (Grob et al., 2017; Rahman et al., 2021; Weiss and Belland, 2016). Therefore, the purpose of this research was to see how a technology-integrated "planned formative assessment strategy affects retention of students" conceptual and procedural knowledge in learning chemistry in general and chemical equilibrium in particular. To address the above objective, the researchers came up with two specific research questions:

1. When comparing time points and intervention groups what, if any, change is there in students' conceptual knowledge examination scores?
2. Is there a difference in students' procedural test scores when time points and intervention groups are compared?

Review of Related Literature

Educators and educational experts have been attempting to integrate technology into their classrooms for the past three decades (Keengwe et al., 2012). It entails the use of technology resources such as computers, mobile devices such as smartphones and tablets, digital cameras, social media platforms, software applications, and the Internet in regular classroom activities and school management. When students can properly integrate technology, they may choose technological tools to help them acquire information in a timely manner, evaluate and synthesize the content, and communicate it professionally (Wang et al., 2014). Technology should be as readily available as all other teaching materials, and it should become an integral component of how the classroom functions (Davies et al., 2013). Technology is also transforming the way teachers teach, allowing them to reach out to a wider range of students and assess their understanding in a number of ways (Davies et al., 2013). Technological practice not only enhances students' learning and understanding, but it also increases their passion for studying, which is crucial for learning, stimulates collaborative learning, and aids in the development of problem-solving skills (Blau et al., 2020).

Incorporating technology into the classroom, according to educational scholars, may assist both students and teachers. For example, technology may help learners become more motivated and provide them with important skills that will help them reinforce their learning (Mayer, 2011). To properly incorporate technology, learners must learn to use the technical instruments

Science Education International
33(4), 383-391
https://doi.org/10.33828/sei.v33.i4.5

necessary for information retrieval. As a result, technology in educational settings, such as other classroom tools, should be available. The teacher who effectively incorporates technology into the classroom considers how to use tools in the classroom to encourage learning without prompting rather than how to use them to use them to promote learning. As a result, although technology facilitates the acquisition of content knowledge, it also facilitates the acquisition of content knowledge.

Smart phones, computers, the Internet, classroom smart boards, and social media are all examples of technology-rich environments that have had a big influence on education. Advances in technology, as well as the increased usage of technology-rich environments in educational settings, have generated discussions regarding assessment in the digital age and technology-enhanced assessment (Timmis et al., 2016). Due to a lack of exact and recognized definitions, theoretical multidimensionality of the constructs, and subjectivity in scoring, evaluating complicated skills (e.g., creativity, problem solving, and critical thinking) was challenging in the past (Shute et al., 2016).

Technology's role in helping formative assessment methods is widely understood. Since technology such as mobile devices, computers, tablets, and online resources is becoming more widely used in classrooms, researchers are interested in examining and analyzing the effectiveness of using digital technologies for formative assessment (Gikandi et al., 2011). According to Aldon and Panero (2020), technology may assist students and facilitators by offering a formative assessment of skills, knowledge, and understanding. Computers might be used to aid in the study of the subject and issue, with the facilitator connecting to liquid crystal display (LCD) projectors to convey information to the students. Because computers make it easy to provide mark lists through spreadsheets, information, and activities, assessment and feedback should no longer be a problem. Problem-solving, presentations, report writing, online research, and receiving tutorials are all examples of assessment assignments that students might do electronically. One of the key limitations to improving formative assessment in the past, according to Redecker and Johannessen (2013) and others, was that the vast majority of courses was delivered face-to-face, preventing the capture of learning interactions and outcomes into a system for identifying and analyzing formative feedback and assessment. Frequent formative assessments and comments for a large group of students can take up a lot of time for teachers, which can be frustrating in practice (Burns et al., 2010). This is one area where technology may be really useful. In both face-to-face and online sessions, students benefit from technology-integrated formative assessment since it provides them with meaningful feedback and personalizes their learning experience (Matuk et al., 2015).

The primary goal of technology-integrated formative assessment in the classroom is to offer students with relevant and timely feedback, and the second goal is to assist students in personalizing their learning. As a result, enhancing the quality of feedback and the manner in which it is presented are crucial components of learning and technology can assist in achieving these two goals (Timmis et al., 2016). In the 21st century, there are several options to record and analyze both performance and assessment data using modern technology to understand how students grow with various sorts of activities and then regulate what alterations may be made to assist individual learners. Critical thinking and problem solving are two skills that new technology may help with in the 21st century. As a result, new assessment approaches are required as a result of the rising use of technology in teaching and learning. According to Shirley and Irving (2015), technology facilitates formative assessment by providing immediate feedback to both learners and teachers and encouraging student participation. As a result, it is possible that more focused instruction and engaged learning may improve.

The teacher's involvement is critical when it comes to using technology in the classroom, as the teacher's approach has an influence on the educational function of technology (Shernoff et al., 2017). Resta and Laferrière (2007) stated in their study of a networked system that teachers need clear pedagogical patterns or teaching routines to effectively engage students in collaborative learning through digital technology. Despite the fact that their teaching techniques recommendations are context-specific, they demonstrate how learning may be improved through a formative process affected by technology. However, opinions disagree on whether technology lowers the teacher's role to that of a "sidekick" or whether technological techniques are only a supplement to teachers rather than a replacement (Jeffrey et al., 2014). Both viewpoints imply that pedagogical adjustments are required, but the scope and character of the changes are unclear.

One of the limiting factors in the development of technologically advanced educational systems appears to be the difficulty in implementing pedagogical modifications (Ertmer and Ottenbreir-Leftwich, 2013). To effectively use technology in learning processes, teachers must employ appropriate pedagogical strategies; however, focusing just on enhancing teachers' digital technology competence while disregarding associated pedagogical consequences looks unlikely to be adequate (Ertmer et al., 2012). This fundamental knowledge includes an understanding of how to communicate concepts and pedagogy while using technology, as well as an awareness of how technology may help students with their conceptual issues. Almalki and Gruba (2020) recommend that teachers' understanding of hardware and software be supplemented by a grasp of formative assessment and related pedagogies in their research on technologically enhanced formative assessment.

RESEARCH METHODOLOGY

Research Design

The quasi-experimental research design was employed in this study, which included a non-equivalent pretest, multiple treatments, and a post-test control group. With pre-test,

Science Education International
33(4), 383-391
https://doi.org/10.33828/sei.v33.i4.5

post-test, and follow-up post-test, the design contains one comparison group and two experimental groups. As a result, the research design for this study may be summarized as follows: Group 1 received Technology-integrated Formative Assessment (TIFA), Group 2 received Formative Assessment (FA) alone, and Group 3 received conventional methods (see Table 1).

Population and Sampling Technique

The present study focused on secondary education, namely, second cycle secondary schools, in Addis Ababa, Ethiopia. As a result, the study's participants were grade 11 students (18–20-years-old) in government secondary schools. The Addis Ababa Administration City was chosen as a convenient sample location for the study. Out of ten Addis Ababa sub-cities, the lottery method was used to choose three as the target population. Three schools from each of the three sub-cities were chosen as a sample using a random sampling method. Three sections from each of the three schools were chosen at random using a sampling technique. Three sections of the schools were randomly assigned, two experimental and one comparison, using a random sampling method. Then, for each school, one chemistry teacher was purposefully chosen who was relatively well qualified and experienced in teaching chemistry.

Variables of the Study

The intervention groups and time point were the study's independent variables. There are three levels of intervention groups. The two experimental groups were given just the independent variables TIFA and FA, whereas the comparison group was given the same conditions. There are three levels of time points as well (pre-test, post-test, and follow-up test). Students' conceptual and procedural knowledge scores from chemical equilibrium examinations were the study's dependent variables.

Instruments for Data Collection

The chemical equilibrium conceptual test (CECT) and chemical equilibrium procedural test (CEPT) were utilized as data collection techniques in the study to answer the research questions. To reduce students' remembered effects on test topics, the researcher created pre- and post-test conceptual and procedural questions that were similar but not identical. All questions were adapted from literature relevant to chemical equilibrium and modified to fit the study's objectives to measure the students' learning outcomes in conceptual and procedural knowledge. The reliability coefficient calculated for internal consistency of all conceptual and procedural test items was 0.74 and above, which was within the acceptable range (Mensah, 2017; Özmen, 2008)**.**

Validity and Reliability of the Instruments

The instrument's items were a collection of questions published by other researchers. The questions covered every aspect of the chemical equilibrium syllabus's erroneous themes. The chemical equilibrium conceptual and procedural examinations were examined for both content and face validity. Instruments for data collection: Experts in chemistry were given pre-conceptual tests, post-conceptual tests, pre-procedural tests, and post-procedural examinations. PhD candidates in chemistry education and secondary school chemistry teachers assessed the test items for compatibility with the textbook objectives and the items, as well as for clarity and errors in the answer key. Finally, the expert's opinions and recommendations were taken into account while making adjustments.

Furthermore, the researchers believed that estimating the internal consistency dependability of quantitative research instruments during the pilot test was adequate to verify the instrument's reliability. The study was conducted at one school that was not included in the study sample of 40 students in 12th grade. The pre- and post-tests for conceptual and procedural chemical equilibrium were piloted with students' who volunteered to help with the instrument and research design. The Kuder-Richardson formula 20 (K-R20) was used to obtain a reliability coefficient estimate of roughly 0.72 for CECTs and 0.75 for CEPTs, respectively.

Procedures for the Treatments

The study included three groups: Two experimental and one comparison. Before receiving treatment, the pre-test chemical equilibrium conceptual and procedural tests were administered to each group from the comparison and experimental categories. On the one hand, the formative assessment alone group was exposed to planned for interaction formative assessment activities that aim to develop conceptual and procedural knowledge through the use of a variety of examples of conceptual and procedural problems, but every activity in and out of the classroom was delivered without supporting technology. The technology-integrated planned formative assessment group, on the other hand, received the same content while being explicitly introduced to technology-supported discourse that incorporates the three components of macro-micro-symbolic teaching as well as every formative activity supported by technological tools and software over the course of the study.

A computer desktop, plasma screen, a laptop, a white board, a microphone, and a smart phone were among the technical equipment employed in this investigation. Telegram, PowerPoint, and internet access were among the programs utilized. The goal of employing such technical tools and software was to make formative assessment procedures easier to use both within and outside the classroom. The teacher used Power Point to develop individual and peer formative activities as well as the course objectives and success criteria. Individual and peer formative activities were two different types of inquiries. The teacher introduced the lesson goals in the classroom by presenting the activities on a computer desktop with the help of plasma screen. During this period, the teacher allows enough time for individual and group discussions about the formative tasks.

The teacher's role in this classroom was to help and guide the students. In addition, the teacher exhibited the scientific answers on the plasma screen after giving the formative tasks.

In addition, the teacher and the students created a telegram group. This telegraph group's work displays their usage of telegrams, and the teacher always attached both conceptual and procedural assignments to them so that students could complete them at home. The teacher utilized a telegram to convey blame outside of the classroom whenever a student made a mistake. The teacher also used this telegram group to link the necessary instructional resources, helping the students develop their conceptual and procedural comprehension.

To encourage discussion amongst the students, the lessons were usually performed through individual and cooperative group work. Throughout the session, students were asked questions that tested high-level thinking abilities and encouraged them to ponder, and they were given time to consider before responding. To achieve so, teachers used the formative assessment techniques concept map, conceptual diagnosis, observation, self-assessment, quiz, oral questions, think-pair-share, one question and one comment, 3-min pause, and 1-min essay in the classroom. This means that when teachers use formative assessment strategies to teach for the advancement of students' higher-order cognitive knowledge, they must offer meaningful feedback during each task. On the other hand, in the comparison group, the teacher used the existing instruction. The conceptual and procedural knowledge examinations were conducted as a post-test after the research period was ended, and the conceptual and procedural knowledge scores were compared. Finally, follow-up post-test was administered for all groups to evaluate the time point effects of each group after 2 months.

Methods of Data Analysis

The results obtained from all the instruments administered were coded and analyzed by the researchers. To make the analyses, more valid and reliable mixed model analysis of variance (ANOVA) was conducted. Before the analysis, the needed assumptions were investigated for testing. In this way, univariate and multivariate normality, homogeneity of variances, sphericity, and variance-covariance homogeneity assumptions were analyzed (Tabachnick and Fidell, 2014). This analysis was done with the help of Statistical Package for Social Sciences (SPSS) computer package version 26.

Table 1: The diagrammatic representations of non-equivalent comparison group research design

Groups	Pre-test	Treatments	Post-test	Follow-up post-test
Treatment Group 1 (TIFA)	O_1	E_1	O_2	O_3
Treatment Group 2 (FA)	O_1	E_2	O_2	O_3
Comparison Group (CG)	O_1	X	O_2	O_3

Creswell (2014)

Table 2: Means and standard deviations for the conceptual knowledge test scores as a function of a 3 (Time) × 3 (Groups)

Time	Group	M	SD	*n*
Pre-conceptual Knowledge test	TIFA	7.87	2.64	45
	FA	6.95	3.08	43
	CM	8.27	2.49	44
	Total	7.70	2.78	132
Post-conceptual knowledge test	TIFA	18.93	2.30	45
	FA	16.63	3.86	43
	CM	15.05	4.06	44
	Total	16.89	3.97	132
Follow-up conceptual knowledge test	TIFA	17.27	2.74	45
	FA	14.74	3.30	43
	CM	12.57	3.13	44
	Total	14.88	3.60	132

Table 3: Mixed Model ANOVA Results for Time Point by Groups on Conceptual Test Scores

Source	Type III MS	Df	MS	F	ρ*	η^2
Between-subject effective						
Group	524.07	2	262.03	16.73	0.000	0.21
Error	2020.23	129	15.66			
Within-subject effect						
Time	6133.43	2	3066.72	419.04	0.000	0.77
time*group	348.39	4	87.10	11.90	0.000	0.16
Error (time)	1888.15	258	7.32			

*0.05. ANOVA: Analysis of variance

MAIN FINDINGS

Students' Conceptual Knowledge and the Effects of Time Points by Groups

A two-way, 3 (time: Pre-test, post-test, and follow-up test) × 3 (groups: TIFA, FA alone, and CM groups) mixed ANOVA with repeated measures on conceptual knowledge test scores was conducted. The results of the descriptive and inferential statistics are presented in the following Tables 2 and 3.

The two-way mixed ANOVA findings demonstrated a significant main effect across groups on conceptual test scores with a large effect size ($F (2, 129) = 16.73$, $\rho < 0.001$, $\eta_p^2 = 0.21$). Furthermore, there was a significant main impact of time point on conceptual test scores ($F (2, 258) = 419.04$, $\rho < 0.001$, $\eta_p^2 = 0.77$) with a large effect size. Furthermore, on conceptual knowledge test scores, there was a significant interaction between time point and group ($F (4, 258) = 11.59$, $\rho < 0.001$, $\eta_p^2 = 0.16$) (Table 3). Examination of the cell means indicated that there was a large increase in conceptual test scores from per-test (*M* = 7.70, SD = 2.78) to post-test (*M* = 16.89, SD = 3.97). However, there was a decrement in conceptual test scores form post-test (*M* = 16.89, SD = 3.97) to follow-up test (*M* = 14.88, SD = 3.60) (Table 2). The profile plots, Figure 1, illustrate the trend across time for three intervention groups. The tread was rising in a linear pattern, and the three groups' tendencies were similar. Overall, the three groups differed, with the conventional approach group

Science Education International
33(4), 383-391
https://doi.org/10.33828/sei.v33.i4.5

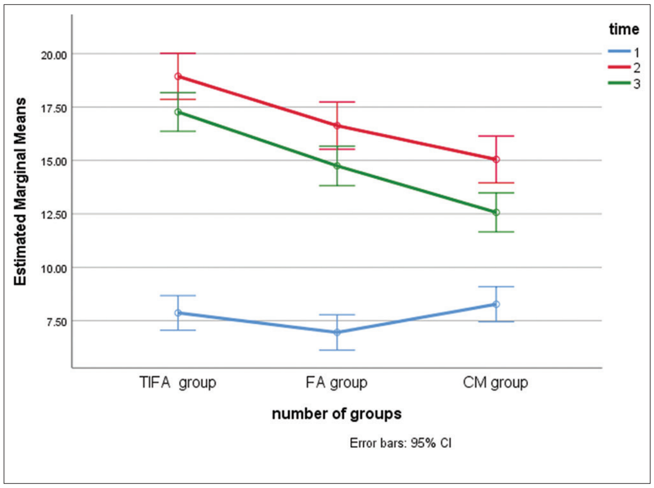

Figure 1: Line Plot Represents Students' Conceptual Knowledge Test Scores across Different Time Point among Groups

being less effective than the other two. Furthermore, the three groups' mean improvements in post-test were larger than the test scores at the other 2-time periods.

The Influence of Time Points by Groups on Students' Procedural Knowledge

A two-way, 3 (time: Pre-test, post-test, follow-up test) × 3 (groups: TIFA, FA alone, CM groups) mixed ANOVA with repeated measures on procedural knowledge test scores was also conducted. Mauchly's test for procedural test indicated that the assumption of sphericity had been violated, W = (χ^2(2) = 16.8, $\rho < 0.05$); therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = 0.65$). Each descriptive and inferential statistical outcome is examined in detail in Tables 4 and 5.

There was a significant difference across the three time points, F (1.30, 167.97) = 773.19, $\rho < 0.001$ and significant differences among groups, F (2, 129) = 24.03, $\rho < 0.001$, in procedural test scores. According to Cohen (1992) guidelines, the effect size of both groups and time points was large. F (2.60, 167.92) = 9.18, $\rho < 0.001$, was also found to be a significant interaction between time point and group. However, the effect size for the interaction between time points and groups was medium. Following up this interaction, it was indicated that there was no significant mean difference among groups at baseline as compared to the 2-time points. Looking at the descriptive data, we can observe that the mean gain in procedural knowledge test scores increased dramatically from pretest (M = 3.78, SD = 2.07) to posttest (M = 10.39, SD = 1.68). On the other hand, from post-test (M = 10.39, SD = 1.68) to follow-up test (M = 9.23, SD = 1.86), there was a decrease in mean gain in procedural knowledge test scores (Table 4). Furthermore, we can observe from the profile plot of means that the mean post-test procedural knowledge test scores were greater than the mean follow-up and pre-test procedural knowledge test scores. At each time point, however, the mean gain of the TIFA group was larger than the mean increase of the FA alone and CM groups (Figure 2).

Table 4: Means and standard deviations for the procedural knowledge test scores as a function of a 3 (time) × 3 (Groups)

Time	Group	M	SD	*n*
Pre-procedural knowledge test	TIFA	4.09	1.62	45
	FA only	3.39	2.52	43
	CM	3.84	1.96	44
	Total	3.78	2.07	132
Post-procedural knowledge test	TIFA	11.56	1.49	45
	FA only	10.12	1.45	43
	CM	9.45	1.37	44
	Total	10.39	1.68	132
Follow-up procedural knowledge test	TIFA	10.64	1.42	45
	FA only	9.12	1.48	43
	CM	7.89	1.56	44
	Total	9.23	1.86	132

Table 5: Mixed Model ANOVA Result for Time Point by Groups on Procedural Knowledge Test Scores

Source	Type III MS	Df	MS	F	ρ*	η^2
Between-subject effective						
Group	205.89	2	102.95	24.03	0.000	0.27
Error	552.61	129	4.28			
Within-subject effect						
Time	3277.36	1.30	2517.72	773.19	0.000	0.86
Time* group	77.80	2.60	29.89	9.18	0.000	0.13
Error (time)	546.80	167.92	3.26			

*0.05

DISCUSSION AND CONCLUSIONS

Discussion of Main Findings

The main and interaction effects of time point versus group differences as repeated measurements of students' conceptual and procedural test scores were evaluated using a mixed-model factorial ANOVA. A follow-up post-chemical equilibrium conceptual and procedural knowledge test was used to determine the mean gain of concepts. According to descriptive statistics (Tables 3 and 5), the TIFA group outperformed the FA-alone group and the CM group on conceptual test scores. There was a difference in effectiveness between the three groups, with the conventional method group being less effective than the other two. Furthermore, in the post-test, the three groups' mean improvements were greater than in the other 2-time periods (Figure 1). Similarly, the mean scores on the post-test procedural knowledge examination were higher than the mean scores on the follow-up and pre-test procedural knowledge test. At each time point, however, the mean gain of the TIFA group was larger than the mean increase of the FA alone and CM groups (Figure 2).

Science Education International
33(4), 383-391
https://doi.org/10.33828/sei.v33.i4.5

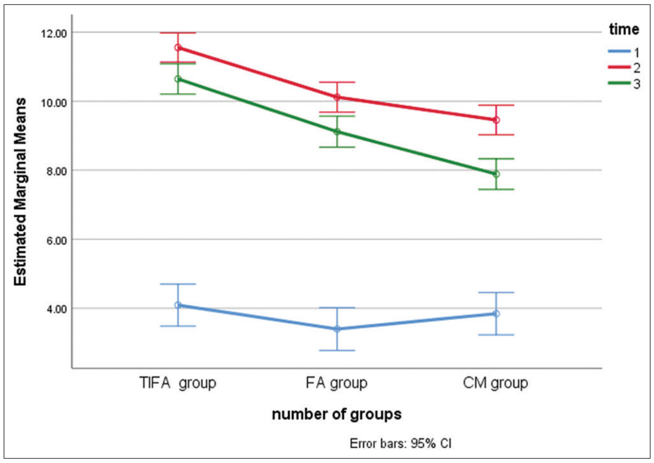

Figure 2: Profile Plot Represents Students' Procedural Knowledge Test Scores across Different Time Point among Groups

The three groups had varied mean gain scores in conceptual and procedural examination scores as repeated measurements of time, as can be observed from the descriptive statistics. A two-way mixed ANOVA was used to see if the differences were statistically significant. The two-way mixed ANOVA findings revealed that there was a significant main effect between groups (F (2, 129) = 16.73, $\rho < 0.001$, $\eta_p^2 = 0.21$) on conceptual test scores. Furthermore, on conceptual knowledge test scores, there was a significant interaction (F (4, 258) = 11.59, $\rho < 0.001$, $\eta_p^2 = 0.16$) between time point and intervention groups (Table 3). On procedural examination scores, there was also a significant difference between the 3-time points (F (1.30, 167.97) = 773.19, $\rho < 0.001$, $\eta_p^2 = 0.86$) and intervention groups (F (2, 129) = 24.03, $\rho < 0.001$, $\eta_p^2 = 0.27$). A significant interaction (F (2.60, 167.92) = 9.18, $\rho < 0.001$, $\eta_p^2 = 0.13$) between time points and groups was also observed (Table 5).

According to Cohen's (1992) recommendations, the impact size of both groups and time points was large in both conceptual and procedural test scores (Tables 3 and 5). The interaction between time points and groups had a large impact size, as shown in Tables 3 and 5. In comparison to the control group, students in the experimental groups demonstrated a considerably deeper conceptual and procedural knowledge of the unit's contents and were better able to explain their thoughts in a clear way. This suggests that formative classroom assessment strategies might help students improve their learning results.

This conclusion backs up the Matilda and Helen (2019) findings, which revealed that formative feedback, had a substantial impact on junior secondary school basic science students' success in the experimental group rather than the control group in their study. Furthermore, the findings are consistent with those of Shah (2014), who found that the experimental group's mean performance scores on a post-test were considerably higher than the control group. This is not unexpected because, according to Alashwal (2020), formative assessment is beneficial to students because it assists in the diagnosis of learning challenges and the understanding of topic difficulties, allowing students to enhance their academic learning outcomes.

The study's findings also revealed that a technology-integrated formative assessment strategy is effective in improving secondary school students' conceptual and procedural test scores, as it was discovered that there was a significant difference between the pre-test, post-test, and follow-up post-test scores in the experimental group, with the difference favoring the post-test scores. This is not unexpected, given that research has shown that students who receive such feedback become more attentive and engaged in the learning process, and they begin to perceive the evaluation process as a tool to help them progress. As a consequence, when students responded to assessment findings by understanding where they were on the path to success, where they were going, and what they needed to do to get there, they tended to learn more and succeed at a higher level (Agbatogun, 2013; Rinaldi et al., 2017).

In general, while there is data that show that employing technology has beneficial advantages, it is difficult to make firm conclusions on its usage in classrooms. While some studies support assertions that technology can facilitate deeper learning (VanderArk and Schneider, 2012) and offer evidence of learning advantages (Higgins et al., 2012), other studies are less conclusive (Haßler et al., 2016) and contradict these research findings. Some of these differences in effect could be due to differences in how technology is used (Higgins et al., 2012) or the pedagogy used by teachers (Fullan and Donnelly, 2013) but there is a clear need to better understand the interactions between teacher, student, and technology and how they can support formative assessment.

CONCLUSION

The technology-integrated formative assessment strategy is successful in raising secondary school students' conceptual and procedural test scores. This was confirmed by the finding that there was a significant difference between the pre-test, post-test, and follow-up post-test scores in the experimental group, with the difference favoring the post-test scores. The most evident outcome of this research is that adopting technology-based tools and software increases formative assessment and, as a result, student learning.

CONFLICTS OF INTEREST STATEMENT

The authors state that there were no commercial or financial ties that may be considered as a possible conflict of interest during the research.

ETHICAL STATEMENT

In this study, the researchers considered the following ethical issues during data collection, interpretation and dissemination (Addis Ababa University). First, the researcher sought permission from the school administration to allow the researchers to conduct the study. Second, the researchers

Science Education International
33(4), 383-391
https://doi.org/10.33828/sei.v33.i4.5

were discussed the objectives of the study with the research participants and obtained their informed consent. Third, the researcher promised to behave confidentially regarding the data collected from research participants.

ACKNOWLEDGMENTS

We thank the students and teachers, who participated in this study. We also wish to thank many others for their assistance, support, and guidance.

REFERENCES

Adesina, A.O. (2016). *A Semi-Automatic Computer-Aided Assessment Framework for Primary Mathematics*. (Doctoral Dissertation). United Kingdom: Loughborough University.

Agbatogun, A.O. (2013). *Comparative Effect of Interactive Mobiles (Clickers) and Communicative Approach on the Learning Outcomes of the Educationally Disadvantaged Nigerian Pupils in ESL Classrooms.* Available from: https://hdl.handle.net/1842/7932 [Last accessed on 2022 Nov 04].

Alashwal, M. (2020). Effect of formative classroom assessment in Saudi Arabian elementary schools. *Journal of Education and Practice*, 4(1), 18-27.

Aldon, G., & Panero, M. (2020). Can digital technology change the way mathematics skills are assessed? *ZDM Mathematics Education,* 52, 1333-1348.

Almalki, M.S., & Gruba, P. (2020). Conceptualizing formative blended assessment (FBA) in Saudi EFL. In: Yilan, S.M., & Koruyan, K., (Eds.), *ICT-Based Assessment, Methods, and Programs in Tertiary Education.* United States: IGI Global. pp. 65-82.

Anderson, L.W., & Krathwohl, D.R. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives.* London, United Kingdom: Longman.

Barak, M. (2017). Science teacher education in the twenty-first century: A pedagogical framework for technology-integrated social constructivism. *Research in Science Education*, 47(2), 283-303.

Bernal-Ballen, A., & Ladino-Ospina, Y. (2019). Assessment: A suggested strategy for learning chemical equilibrium. *Education Sciences*, 9(3), 174.

Bhagat, K.K., & Spector, J.M. (2017). Formative assessment in complex problem-solving domains: The emerging role of assessment technologies. *Journal of Educational Technology andg Society*, 20(4), 312-317.

Blau, I., Shamir-Inbal, T., & Avdiel, O. (2020). How does the pedagogical design of a technology-enhanced collaborative academic course promote digital literacies, self-regulation, and perceived learning of students? *The Internet and Higher Education*, 45, 100722.

Burns, M.K., Klingbeil, D.A., & Ysseldyke, J. (2010). The effects of technology-enhanced formative evaluation on student performance on state accountability math tests. *Psychology in the Schools*, 47(6), 582-591.

Chen, S.Y., Tsai, J.C., Liu, S.Y., & Chang, C.Y. (2021). The effect of a scientific board game on improving creative problem solving skills. *Thinking Skills and Creativity*, 41, 100921.

Creswell, J. W. (2014). *Qualitative, Quantitative and Mixed Methods Approaches*. California: SAGE Publications

Cohen, J. (1992). Quantitative methods in psychology: A power primer. *Psychological Bulletin*, 112(1), 155-159.

Davies, R.S., Dean, D.L., & Ball, N. (2013). Flipping the classroom and instructional technology integration in a college-level information systems spreadsheet course. *Educational Technology Research and Development*, 61(4), 563-580.

Elmahdi, I., Al-Hattami, A., & Fawzi, H. (2018). Using technology for formative assessment to improve students' learning. *Turkish Online Journal of Educational Technology TOJET*, 17(2), 182-188.

Ertmer, P.A., & Ottenbreit-Leftwich, A. (2013). Removing obstacles to the pedagogical changes required by Jonassen's vision of authentic technology-enabled learning. *Computers and Education*, 64, 175-182.

Ertmer, P.A., Ottenbreit-Leftwich, A.T., Sadik, O., Sendurur, E., & Sendurur, P. (2012). Teacher beliefs and technology integration practices: A critical relationship. *Computers and Education*, 59(2), 423-435.

Feldman, A., & Capobianco, B.M. (2008). Teacher learning of technology enhanced formative assessment. *Journal of Science Education and Technology*, 17(1), 82-99.

Fullan, M., & Donnelly, K. (2013). *Alive in the Swamp: Assessing Digital Innovations in Education*. United Kingdom: Nesta.

Furtak, E.M., Kiemer, K., Circi, R.K., Swanson, R., de León, V., Morrison, D., & Heredia, S.C. (2016). Teachers' formative assessment abilities and their relationship to student learning: Findings from a four-year intervention study. *Instructional Science*, 44(3), 267-291.

Gabel, D. (1999). Improving teaching and learning through chemistry education research: A look to the future. *Journal of Chemical Education*, 76(4), 548.

Gerard, L., Matuk, C., & Linn, M.C. (2016). Technology as inquiry teaching partner. *Journal of Science Teacher Education*, 27(1), 1-9.

Gikandi, J.W., Morrow, D., & Davis, N.E. (2011). Online formative assessment in higher education: A review of the literature. *Computers and Education*, 57(4), 2333-2351.

Gobert, J.D., Sao Pedro, M., Raziuddin, J., & Baker, R.S. (2013). From log files to assessment metrics: Measuring students' science inquiry skills using educational data mining. *Journal of the Learning Sciences*, 22(4), 521-563.

Grob, R., Holmeier, M., & Labudde, P. (2017). Formative assessment to support students' competences in inquiry-based science education. *Interdisciplinary Journal of Problem Based Learning*, 11(2), 1-9.

Habiddin, H., & Page, E.M. (2021). Examining students' ability to solve algorithmic and pictorial style questions in chemical kinetics. *International Journal of Science and Mathematics Education*, 19(1), 65-85.

Haßler, B., Major, L., & Hennessy, S. (2016). Tablet use in schools: A critical review of the evidence for learning outcomes. *Journal of Computer Assisted Learning*, 32(2), 139-156.

Hickey, D.T., Taasoobshirazi, G., & Cross, D. (2012). Assessment as learning: Enhancing discourse, understanding, and achievement in innovative science curricula. *Journal of Research in Science Teaching*, 49(10), 1240-1270.

Higgins, S., Xiao, Z., & Katsipataki, M. (2012). *The Impact of Digital Technology on Learning: A Summary for the Education Endowment Foundation. Full Report.* (ED612174). United States: ERIC. Available from: https://eric.ed.gov/?id=ED612174 [Last accessed on 2022 Nov 04].

Hsu, Y.Y., & Lin, C.H. (2020). Evaluating the effectiveness of a preservice teacher technology training module incorporating SQD strategies. *International Journal of Educational Technology in Higher Education*, 17(1), 31.

Jahnke, I., Meinke-Kroll, M., Todd, M., & Nolte, A. (2020). Exploring artifact-generated learning with digital technologies: Advancing active learning with co-design in higher education across disciplines. *Technology, Knowledge and Learning*, 27, 335-364.

Jeffrey, L.M., Milne, J., Suddaby, G., & Higgins, A. (2014). Blended learning: How teachers balance the blend of online and classroom components. *Journal of Information Technology Education*, 13, 121-140.

Kay, R.H., & Knaack, L. (2009). Assessing learning, quality and engagement in learning objects: The Learning Object Evaluation Scale for Students (LOES-S). *Educational Technology Research and Development*, 57(2), 147-168.

Keengwe, J., Schnellert, G., & Mills, C. (2012). Laptop initiative: Impact on instructional technology integration and student learning. *Education and Information Technologies*, 17(2), 137-146.

Looney, J. (2019). *Thinking Deeper Research Paper No. 1-Part 3. Making it Happen: Formative Assessment and Educational Technologies*. United Kingdom: Promethean Educational Strategy Group.

Lu, C., & Cutumisu, M. (2021). *Integrating Deep Learning into an Automated Feedback Generation System for Automated Essay Scoring*. In: Proceedings of The 14th International Conference Educational Data Mining Society.

Manuguerra, M., & Petocz, P. (2011). Promoting student engagement by integrating new technology into tertiary education: The role of the iPad.

Science Education International
33(4), 383-391
https://doi.org/10.33828/sei.v33.i4.5

Asian Social Science, 7(11), 61-65.

Matilda, U.O., & Helen, A.O. (2019). Effect of formative classroom assessment on students' academic achievement in junior secondary school basic science in Egor local government area of Edo state. *Journal of Education and Practice*, 10(15), 184-189.

Matuk, C.F., Linn, M.C., & Eylon, B.S. (2015). Technology to support teachers using evidence from student work to customize technology-enhanced inquiry units. *Instructional Science*, 43(2), 229-257.

Mayer, R.E. (2011). Towards a science of motivated learning in technology-supported environments. *Educational Technology Research and Development*, 59(2), 301-308.

Mensah, A. (2017). *Promoting Conceptual Change in Chemical Equilibrium Through Metacognition Development: Students' Achievement and Metacognitive Skills.* (Doctoral Dissertation). Potchefstroom Campus, South Africa: North-West University.

Nakiboglu, C. (2008). Using word associations for assessing non major science students' knowledge structure before and after general chemistry instruction: The case of atomic structure. *Chemistry Education Research and Practice*, 9(4), 309-322.

Özmen, H. (2008). Determination of students' alternative conceptions about chemical equilibrium: A review of research and the case of Turkey. *Chemistry Education Research and Practice*, 9(3), 225-233.

Paiva, S., Reis, L.P., & Raquel, L. (2020). Formative assessment and digital tools in a school context. In: Rocha, A., Adeli, H., Resi, L.P., Costanzo, S., Orovic, I., & Moreira, F., (Eds.), *Trends and Innovations in Information Systems and Technologies,* Vol. 3. Germany: Springer. pp. 271-283.

Quellmalz, E.S., Davenport, J.L., Timms, M.J., DeBoer, G.E., Jordan, K.A., Huang, C.W., & Buckley, B.C. (2013). Next-generation environments for assessing and promoting complex science learning. *Journal of Educational Psychology*, 105(4), 1100-1114.

Quílez, J. (2004). Changes in concentration and in partial pressure in chemical equilibria: Students' and teachers' misunderstandings. *Chemistry Education Research and Practice*, 5(3), 281-300.

Rahman, K., Hasan, M., Namaziandost, E., & Ibna Seraj, P.M. (2021). Implementing a formative assessment model at the secondary schools: Attitudes and challenges. *Language Testing in Asia*, 11(1), 1-18.

Redecker, C., & Johannessen, Ø. (2013). Changing assessment-towards a new assessment paradigm using ICT. *European Journal of Education*, 48(1), 79-96.

Resta, P., & Laferrière, T. (2007). Technology in support of collaborative learning. *Educational Psychology Review*, 19(1), 65-83.

Rinaldi, V.D., Lorr, N.A., & Williams, K. (2017). Evaluating a technology supported interactive response system during the laboratory section of a histology course. *Anatomical Sciences Education*, 10(4), 328-338.

Robertson, S.N., Humphrey, S.M., & Steele, J.P. (2019). Using technology tools for formative assessments. *Journal of Educators Online*, 16(2), n2.

Shah, I. (2014). The role of formative assessment in learning chemistry. *Pakistan Journal of Education*, 31(2), 123-133.

Shernoff, D.J., Sinha, S., Bressler, D.M., & Ginsburg, L. (2017). Assessing teacher education and professional development needs for the implementation of integrated approaches to STEM education. *International Journal of STEM Education*, 4(1), 1-16.

Shirley, M.L., & Irving, K.E. (2015). Connected classroom technology facilitates multiple components of formative assessment practice. *Journal of Science Education and Technology*, 24(1), 56-68.

Shute, V.J., Wang, L., Greiff, S., Zhao, W., & Moore, G. (2016). Measuring problem solving skills via stealth assessment in an engaging video game. *Computers in Human Behavior*, 63, 106-117.

Sirhan, G. (2007). Learning difficulties in chemistry: An overview. *Journal of Turkish Science Education.* 4, 2-22.

Spector, J.M., Ifenthaler, D., Sampson, D., Yang, J.L., Mukama, E., Warusavitarana, A., & Gibson, D.C. (2016). Technology enhanced formative assessment for 21st century learning. *Educational Technology and Society,* 19(3), 58-71.

Tabanchnick, B.G., & Fidell, L.S. (2014). *Using Multivariate Statistics*. 6th ed. United Kingdom: Pearson Education, Limited.

Taber, K. (2002). *Chemical Misconceptions: Prevention, Diagnosis and Cure*. Vol. 1. United Kingdom: Royal Society of Chemistry.

Tan, A.L., & Towndrow, P.A. (2009). Catalyzing student-teacher interactions and teacher learning in science practical formative assessment with digital video technology. *Teaching and Teacher Education*, 25(1), 61-67.

Timmis, S., Broadfoot, P., Sutherland, R., & Oldfield, A. (2016). Rethinking assessment in a digital age: Opportunities, challenges and risks. *British Educational Research Journal*, 42(3), 454-476.

Tissenbaum, M., & Slotta, J. (2019). Supporting classroom orchestration with real-time feedback: A role for teacher dashboards and real-time agents. *International Journal of Computer Supported Collaborative Learning*, 14(3), 325-351.

VanderArk, T., & Schneider, C. (2012). *How Digital Learning Contributes to Deeper Learning.* Available from: https://www.worldwideworkshop.com/pdfs/GettingSmart_DigitalLearningDeeperLearning.pdf [Last accessed on 2013 Jun 30].

Wang, S.K., Hsu, H.Y., Reeves, T.C., & Coster, D.C. (2014). Professional development to enhance teachers' practices in using information and communication technologies (ICTs) as cognitive tools: Lessons learned from a design-based research study. *Computers and Education*, 79,101-115.

Wei, L. (2011). Formative assessment in classrooms: Operational procedures. *Journal of Language Teaching and Research,* 2(1), 99-103.

Weiss, D.M., & Belland, B.R. (2016). Transforming schools using project-based learning, performance assessment, and common core standards. *Interdisciplinary Journal of Problem-Based Learning*, 10(2), 1-4.

Wells, G. (2002). Learning and teaching for understanding: The key role of collaborative knowledge building. *Advances in Research on Teaching*, 9, 1-42.