Effects of technology-integrated chemistry instruction on students' academic achievement and retention capacity

Derejaw Yesgat[1], Solomon Melesse[2], Dereje Andargie[3], Belete Bedemo Beyene[4]
[1]Chemical Education, Department of Chemistry, Bahir Dar University, Bahir Dar, Ethiopia
[2]Curriculum and Instruction, Department of Teacher Education and Curriculum Studies, College of Education, Bahir Dar University, Bahir Dar, Ethiopia
[3]Chemistry Education, Department of Teacher Education and Curriculum Studies, College of Education, Debre Berhan University, Debre Birhan, Ethiopia
[4]Organic Chemistry, Department of Chemistry, College of Natural Sciences, Bahir Dar University, Bahir Dar, Ethiopia

Article Info

Article history:

Received Jun 23, 2023
Revised Aug 23, 2023
Accepted Sep 15, 2023

Keywords:

Academic achievement
Chemistry
Instruction
Retention
Technology-integrated

ABSTRACT

This study determined how chemistry instruction using technology affected student achievement and retention. To achieve the goal of the study, Solomon's four-group quasi-experimental research design was used. The one-way analysis of variance (ANOVA) and independent sample t-test were used to statistically examine the data. A one-way ANOVA analysis revealed that, for achievement and retention, respectively, the mean scores of both trial groups showed statistically significant differences from the mean scores of comparison groups $F(3.161)=88.568$, $p<0.05$ and $F(3.161)=14.75$, $p<0.05$). The independent samples t-test on the achievement and retention post-test mean score of the experimental and comparison groups respectively, was statistically significant ($t(163)=16.05$, $p<0.05$ and $t(162.09)=7.61$, $p<0.05$). Independent sample t-test shows that male and female students' post-test scores for achievement or retention in the experimental group results was ($t(42)=1.53$; $p>0.05$ and $t(38)=-0.465$; $p>0.05$, respectively. This intervention has no statistically significant gender related impact on students' achievement and retention. Thus, it is conceivable to draw the conclusion that chemistry instruction that incorporates technology enhances student achievement and retention, but its effect is gender neutral. According to recent studies, chemistry teachers and teacher training programs should use technology-integrated lessons to help students acquire and recall chemical bonding concepts.

Corresponding Author:

Derejaw Yesgat
Chemical Education, Department of Chemistry, Bahir Dar University
H9FX+Q62, Bahir Dar, Ethiopia
Email: yesgatderejaw@gmail.com

1. INTRODUCTION

Students find it challenging and confusing to learn chemistry at the secondary school level since it involves picturing particulate level interaction of atoms, molecules, ions, and indistinguishable and intangible particulate notions [1]. The particle form of matter underlies nearly every problem in chemistry. Thus, it is important for chemistry students to comprehend scientific concepts at the level of the particle of matter, explain the properties of matter, and the various chemical changes that happen in a wide variety of chemical occurrences [1]. Understanding microscopic explanations of how compounds are created and what functions they serve is crucial to students' success in chemistry classes. These microscopic worlds are typically unrelated to the student's daily experiences, making them challenging to understand [2]. Therefore model based-inquiry teaching [3] using technology might enhance learners understanding of microscopic world of chemistry. The use of educational technology as well as the preparation and implementation of technology-based teaching are important factors in improving educational quality, particularly in science education [4], [5]. One method or resource being used to close the current gap and encourage independent and active learning is the use of technology tools. It is critical to realize that computer-assisted technologies help students understand microscopic and sub-microscopic (particulate) topics in chemistry and help them become more prepared to learn. It is important to comprehend particulate matter because it is required to discuss almost every chemical topic, which explains why it is so important [6].

It would seem that the method used to teach the concept of chemistry is essential for concepts to be properly acquired and understood. In an effort to defend student under performance and low retention capacity in science fields, a number of explanations have been put forth. Nwagbo [7] asserts that science topics were not fully understood by students, and that this was evidenced by their persistently poor performance and forgetfulness of science. Therefore, it is essential to choose efficient teaching strategies that will accomplish the purpose of science instruction, enhancing student learning and retention of science courses. Learning is improved for students when technology is incorporated into curriculum and teaching [8], [9]. A web-based computer simulation is used to teach chemistry, Frailich *et al.* [10], on the assumption that it aids students in comprehending the abstract and challenging concepts covered in chemistry classes. Students who use simulations can improve their understanding of chemistry concepts and grow as autonomous thinkers and learners [11]. One method that has received recognition by its potential to teach either challenging or dangerous concepts is computer-assisted instruction (CAI) [12], [13]. It is quite challenging to teach dynamic and complex topics using words, equations, or in-class activities. Computer-based simulation can communicate these concepts by means of a multi-sensory approach that conventional approaches lacks.

Male and female students benefited equally from all instructional approaches, according to Nja *et al.* [14], which means that both genders were equally affected by the web-based computer simulation. Their research found no gender disparities in the learning outcomes from simulated learning environments. It suggests that employing simulation to teach chemical ideas could aid in the elimination of gender stereotypes in science education in this context, as all of the students will enjoy and be driven to learn chemistry. Many studies, including those by Egbodo [15], Mihindo *et al.* [16], and Nkemakolam *et al.* [17], demonstrated that male and female students scored better in chemistry when taught using technology. The aforestated studies indicate that chemistry using technology-integrated instruction has no impact on performance difference between male and female students. On the contrarily, other researchers [18], [19] reported that female students outperformed male students in technology-integrated lessons because they had more processing skill practice. The integrated teaching of science processing skills improves female students more than male students, according to their descriptive design study on computers.

A quasi-experimental research conducted by Tukura *et al.* [20] on the degree of achievement between students who were taught fundamental science topics utilizing e-learning and those who were taught using the lecture technique showed a noticeable degree of achievement difference. The results showed that there was no appreciable difference in performance between students taught using e-learning and those taught using the lecture technique, despite the fact that they were in opposition to earlier survey design research by Tukura *et al.* [20]. Regarding the impact of technology-integrated education on student accomplishment, Carter *et al.* [21] and Carter *et al.* [22] found opposite results. However, their quasi-experimental design study supported the finding that students learning through internet-connected instruction outperformed those taught without internet connection instruction. Their survey design study showed that students learning through web-based instructions performed lower than students learning without web-based instructions.

According to Tukura *et al.* [20], Yusuf and Afolabi [23], and Ani *et al.* [24], no statistically significant performance difference wass observed between male and female students who were taught fundamental science concepts via online education. Of course, according to Mihindo *et al.* [16], Achuonye and Olele [25], male students did superior and gain more from instructional technology than their female equivalents. In contrast, Pate's [26] descriptive survey study on gender difference in achievement during technology-based instruction revealed that females benefited more, and supported the idea that female students are more motivated and have higher expectations of themselves than male students.

There are ample studies (e.g., Mihindo *et al.* [16], Nkemakolam *et al.* [17], Cevahir *et al.* [27], Unal and Yerlikaya [28], Çinar and Çepni [29], Ylmaz and Yanarateş [30], Iyamuremye *et al*. [31], Ardac and Akaygun [32] and Ratamun and Osman [33]) that studied the effect of technology-integrated chemistry teaching on students' achievement. Each one was carried out utilizing a sort of quasi-experimental approach. A descriptive survey study design was utilized by Pate [26] and Yesilyurt *et al.* [34] whereas an experimental research design was used by scholars like Jabeen and Afzal [35], Miller *et al.* [36], and Suleman *et al.* [37]. The fundamental flaw in the descriptive survey and quasi-experimental study designs is that they cannot completely rule out the possibility of confounding bias, which can make establishing causal correlations challenging. This weakness is widely used to cast doubt on the findings of descriptive survey and quasi-experimental research designs. Recent studies that used descriptive survey research designs and quasi-experimental research methods produced conflicting results to support it. There is only one experiment and one comparison group in a quasi-experimental design. With this research design, it is impossible to draw any conclusions about the relationship between an intervention and a particular result [38].

It is commonly known that Solomon's four-group quasi-experimental research design has better advantage to maintain internal and external validity over survey and quasi-experimental designs; Solomon's design rules out a number of internal validity threats (such as history and maturation) and external validity threats (i.e. interactions between pretesting and treatments and between selection bias and treatments. With the help of this design, we can identify when pre-testing effects, also known as test sensitization, occur as well as treatment effects on experimental variables. By analyzing pretest sensitization, Solomon's four groups quasi-experimental method combats the effect of confounding [39].

So far, researchers' have not come across studies conducted on the impact of technology-integrated teaching on secondary school students in Ethiopia's retention and achievement of chemical bonding and structure concepts using Solomon's four-group quasi-experimental designs. Thus, researchers of this study chosen Solomon's four-group design to investigate the impact of technology-integrated education on academic achievement and retention capacity of chemical bonding and structure concepts among Damot Secondary School students. The Solomon design was used to obtain methodologically sound information about technology-integrated instruction. The method, which identifies whether the observed effects were caused by the intervention itself or a range of other factors, has not been used in prior studies in the sector. As a result, this study may contribute to the body of knowledge on technology-integrated instruction by providing insights into the impact of technology integration on Ethiopian secondary school students' understanding of and recall of chemical bonding concepts.

This study was designed to determine the impact of technology-integrated instructions on students' understanding and ability to retain information about chemical bonds. Additionally, it seeks to ascertain the impact of technology-integrated instruction on the disparity in students' academic achievement and capacity to retain information about chemical bonds. The investigation made an effort to address the following research queries: i) Do technology integrated chemistry instruction have a statistically significant effect on students' understanding of chemical bonds?; ii) Do instructions that use technology result in significant gender variations in chemical bonding achievement?; iii) Do lessons that use technology have a statistically significant impact on student's memory of chemical bonds?; and iv) Do technology-integrated instructions result in significant gender disparities in the ability to retain chemical bonds?

2. RESEARCH METHOD

2.1. Research approach

The purpose of the study was to determine how chemistry instruction using technology affected student achievement and retention. Quantitative research methods were applied in this investigation. To examine the effect of technology-integrated chemistry instruction on achievement and retention of chemical bonding knowledge of students in the Damot Secondary School West Gojjam Zone, a quantitative research approach was found suited to the topic under investigation.

2.2. Research design

An experimental study design suitable for intact groups is Solomon's four-group quasi-experimental design with two treatment groups [40]–[42]. In this design researchers can use classrooms exactly as they are set up in educational institutions. This Solomon's group design is frequently employed by medical [43] and educational [44] researchers to explore pretesting effects when randomization into treatment and comparison groups is not feasible. In this Solomon's four-group quasi-experimental design, two experimental groups and two control groups were designed. The number of sampled students is significantly more than in a straightforward quasi-experimental or pure experimental design. In this design if a larger sample size is selected, the average values of the findings are expected to be more accurate. Additionally, a larger sample size reduces the margin of error and aids the identification of data outliers. By analyzing pretest sensitization, Solomon's four groups design also counteract the impact of confounding.

Table 1 shows that the Solomon's design normally consists of two experimental and two control groups, as was the case in this investigation. There were two classrooms in each group. Pre-tests were administered to one group from each of the experimental and comparison groups (two classrooms from each group), but they were not administered to the other group. The experimental and control groups were randomly assigned to the four classrooms. The number of experimental and control group students in the pre-test of this study were 36 and 44 students. Likewise, there were 40 experimental and 45 control group students not in the pre-test category. The instructional interventions lasted 32 classes over the course of eight weeks in experimental classrooms. There were 40 minutes in each class. The chemical bonding achievement test (CBAT) was given to each of the experimental and control groups one week before the intervention started. The CBAT post-test was given twice to each group after one and four weeks of the intervention.

Table 1. Solomon's four-group quasi-experimental design [45]

Group	Pretest score	Intervention	Posttest score
Experimental group 1	O1	Technology-integrated instruction	O2
Comparison group 1	O3	-	O4
Experimental group 2		Technology-integrated instruction	O5
Comparison group 2		-	O6

Note. X: treatment, O: outcomes

2.3. Data collection instruments

CBAT was applied on both the experimental and comparison groups were given the CBAT for the pre- and post-test in order to study the impact of the intervention on academic achievement and retention. The pretest served as the analysis's baseline. The purpose of the post-test is to determine the final score and the variation between the scores of students of both the experimental and control before and after treatment.

2.4. Reliability and validity of CBAT

Eight chemistry instructors from the institution served as the panel of subject matter experts involved to develop CBAT in order to increase its content validity. Before integrating CBAT in the pre- and post-tests, the researchers asked chemical experts for their input to enhance the validity of the CBAT. CBAT test was given for two experienced chemistry teachers who have over 25 years of experience in teaching chemistry to examine its face validity and content validity. The identification of the desired learning outcomes of the curriculum, the focus placed on those outcomes, the preparation of the test and degree of difficulty of the topics was done by experienced school teachers. To further enhance the test's validity, a table of specifications was created and used for CBAT preparation.

2.5. Pilot study

To confirm their reliability, the pre- and post-CBAT were piloted on 50 non-sampled grade 12 students. The grade 12 students who were not part of the sample were used in a pilot study at another secondary school to prevent information from tainting the test results. The researchers chose to test the pilot on grade 12 students because the students who took part in it had studied chemical bonds in grade 11.

2.5.1. Reliability of chemical bonding achievement test

The reliability tested using the Kuder Richardson reliability or KR-20 formula. The test was done, based on the variation of the total scores as well as the percentage of correct and incorrect answers to each test question. For the pre- and post-tests, the reliability was discovered to be 0.853 and 0.811, respectively, as seen in Table 2. Because of this, it can be concluded that the tests were accurate at assessing students' understanding of chemical bonding concepts in the previously mentioned chemistry courses because the reliability test coefficient or reliability index is more than 0.7 for both the pretest and the post-test [46]. The results of the pilot test were to examine the difficulty level (P) and discrimination power (D) of each multiple-choice items on both the pre- and post-tests in addition to the tests' reliability.

Table 2. Reliability statistics of pre-and post-CBAT

	KR-20 reliability test coefficient	N of Items
Pre-CBAT	0.853	20
Post-CBAT	0.811	30

2.6. Data collection procedure

The achievement test on chemical bonding content was given to two of the four groups (i.e. one comparison group and one experimental group) as a pretest in order to examine the impact of technology-integrated instruction on the academic performance and retention of secondary school students in West Gojjam Zone. However, no pretest was administered to the remaining comparison and experimental groups. Lesson plans for all the four groups were prepared after the pretest to teach about chemical bonding. The two of the experimental groups received training with the aid of technology. The same teacher instructed the two comparison group students using a non-technological strategy known as a "talk-and-chalk" style lesson, which is the predominant teaching strategy in Ethiopia in general and so is in the school. The syllabus allocated eight weeks for teaching chemical bonding. The experimentation lasts for eight weeks because the intervention is carried out concurrently with the regular calendar. All the four groups had taken CBAT test as a post-test after a week after concluding the intervention to investigate the impact of technology-integrated instruction on students' academic achievement on chemical bonding and capacity to remember. The CBAT post-test were again administered by conventional classroom teachers as delayed retention post-tests after four-week intervals to determine their retention level, but the original test was switched around.

2.7. Data analysis method

In this study an independent sampling T-tests, one-way and 2x2 factorial analysis of variance (ANOVA) statistical tools were used. The essential assumptions were verified prior to starting the analysis. Analysis was done on the variances' normality and homogeneity. In terms of normality for univariate, skewness and kurtosis coefficients were obtained for each group and variable. Pretest sensitization was investigated using a 2 (treatment, yes or no) × 2 (pretest, yes or no) between-group ANOVA on the four post-test scores. A one-way ANOVA was utilized to determine whether there is a significant difference between the posttest achievement and retention of the four groups. In order to determine whether there is a significant difference in the achievement and retention capacities of boys and females, an independent sampled T-test was utilized. All statistical results were subjected to statistical significance tests at the alpha level of 0.05, to draw reliable inferences from the data.

2.8. Ethical concerns

Ethics issues were considered seriously in this investigation. The researchers had frank discussions with the participants about the study's goals, its advantages, their privacy, and the lack of risks involved in taking part. The researchers received permission and authorization to do their research in schools. Consequently, participants of the study, (i.e., teachers and students) were involved in the study based on their full agreement. The legislative structure of the educational system was used to get the consent. Additionally, a Bahir Dar University review board or ethics board granted ethical clearance and supervised the research.

3. RESULTS

3.1. Analysis of pre-test sensitization

Solomon's four-group design analysis starts by determining whether there is proof of pretest sensitization. The test for this uses a between-groups factorial ANOVA on the four post-test scores with two groups (i.e., pretest or no pretest) x two groups (i.e., usage of treatment or no treatment). In this study, the CBAT post-test scores of the four groups were compared using a 2 (test; pre-tested or no pre-tested) 2 (treatment; experimental or control) factorial ANOVA. ANOVA results were also obtained for the main effects of post-test (experimental and control groups) and testing (pre-tested and no pre-test groups).

As seen from Table 3, a 2-factorial ANOVA table, the main effect of the intervention on post-mean achievement gives F statistics for a group $F(1.161)=258.27$, $p<0.05$. The result shows a significant difference between the treatment group and the comparison group. However, no significance difference was observed between the pre-tested and non-pretested groups ($F(1.161)=1.777$, $p>0.05$). The interaction effect between groups and the pre-test was non-significant ($F(1.161)=2.932$, $p>0.05$. These rows informed us whether the independent variables (the "pretest" and "group" rows) and their interaction (the "pretest*group" row) have a statistically significant effect on students' achievement. As seen from Table 3, no significant interaction difference was observed between the pre-test and the comparison group ($p<0.05$). The interaction effect of pretest × groups at ($p>0.05$) was not significant, but there is a significant difference between groups at ($p\leq0.05$). Hence, the results revealed that pretest sensitization was not regarded as a threat to the internal validity of the current experiment.

As can be seen from Table 4, a 2×2 factorial ANOVA table of retention test scores gives F statistics for a group ($F(1.161)=55.784$, $p<0.05$; when the main result of treatment was taken into account, a significant difference was observed between the experimental and comparison groups. On the other hand, no significance difference was observed between the pre-tested and non-pretested groups ($F(1.161)=0.41$, $p>0.05$). The instruction effect between groups and the pre-test was not significant, $F(1.161)=0.343$, $p>0.05$. The pre-test and group rows inform us whether our independent variables and their interaction (the "Pretest × Group" row) have a statistically significant effect on the dependent variable. There was no significant interaction between pretest × groups at $p>0.05$. However, there is a significant interaction between groups at $p\leq0.05$. Hence, the results show that pretest sensitization could not be not considered as a risk to the internal validity of the current study design.

Table 3. A 2×2 factorial ANOVA analysis of post-CBAT

Source	Type III sum of squares	df	Mean square	F	Sig.	Partial Eta squared
Corrected model	2599.609[a]	3	866.536	88.568	.000	.623
Intercept	66651.945	1	66651.945	6812.43	.000	.977
Groups	2526.931	1	2526.931	258.275	.000	.616
Pretest	17.387	1	17.387	1.777	.184	.011
Groups * pretest	28.683	1	28.683	2.932	.089	.018
Error	1575.203	161	9.784			
Total	69417.000	165				
Corrected total	4174.812	164				

a. R Squared=.623 (Adjusted R squared=.616)

Table 4. A 2×2 factorial ANOVA analysis of retention tests between subjects affects ANOVA results analysis

Source	Type III sum of squares	df	Mean square	F	Sig.
Corrected model	754.942[a]	3	251.647	18.697	.000
Intercept	62477.998	1	62477.998	4641.964	.000
Groups	750.812	1	750.812	55.784	.000
Pretest	5.514	1	5.514	.410	.523
Groups * pretest	4.613	1	4.613	.343	.559
Error	2166.961	161	13.459		
Total	64750.000	165			
Corrected total	2921.903	164			

a. R Squared=.258 (Adjusted R squared=.245)

3.2. Analysis of the pre-test results

The pre- and post-tests results were analyzed based on the study's research questions to examine the effect of technology-integrated instruction on students' achievement and retention capacity. Findings from the baseline data (pre-results analysis) were used in the first part of the analysis section, and the findings from the post-intervention quantitative data were used in the second part of the analysis section. To apply an intervention, assessing participants' baseline achievement is necessary. An independent sample t-test analysis was performed on the mean scores of the pre-tested groups. To determine students' achievement before the intervention, the pre-chemical bonding achievement test was administered to pre-tested groups. It helps to determine whether or not pre-tested groups were equivalent prior to the intervention. An independent sample t-test was conducted to compare the pretest academic achievement scores for the experimental and control groups. As can be shown in Table 5, no statistically significant difference in the students' CBAT scores between the intervention group's performance (M=19.22, SD=5.68) and that of the comparison group (M=19.68, SD=6.13); t (78)=0.345, p>0.05. It should be highlighted that both the treatment and comparison groups' pretest outcomes are comparable. This shows that no statistically significant differences were observed among all study variables before intervention. This suggests that the study's groups shared traits and were thus a good fit for the study.

Table 5. Independent samples t-test of the pre-test mean score on CBAT

	Comparison and experimental groups	N	Mean	SD	Df	t-value	P-Value
Pre-achievement	Experimental group with pretest	36	19.22	5.68	78	-.345	0.73
	Comparison group with pretest	44	19.68	6.130			

Solomon's four-group design has important assumption, which states that the four groups should be comparable in terms of the dependent variable prior to application of the intervention [47]. The equality of the four groups (classrooms) for chemistry achievement was investigated using a one-way ANOVA on their first-semester chemistry final examination scores. Beginning with the pre-test findings, which were utilized for assessing the groups before the intervention baseline knowledge of chemistry topics, this section presents the results in chronological order.

Table 6 presents the four groups equivalence following the grouping step. To see whether the difference was significant, a one-way ANOVA was performed. The one-way ANOVA revealed that the difference in first-semester chemistry final examination achievement scores was not statistically significant (F (3.237)=1.49, p>0.05). First-semester chemistry final examination scores of students in the groups can be taken as equivalent in terms of chemistry achievement, because as it can be seen in Table 6, they had similar chemistry achievements before the intervention.

Table 6. Descriptive statistics of first-semester chemistry final exam achievement

Groups	N	Mean	SD
A	36	40.35	5.45
B	40	38.90	6.42
C	44	41.50	9.34
D	45	39.23	12.9
Total	165	39.99	6.28

3.3. Analysis of the effects of intervention

3.3.1. Analysis of the effects of intervention on students' academic achievement

Post-test results were analyzed using one-way ANOVA to investigate the mean scores differences between the experimental and comparison groups. From Table 7, mean scores of experimental groups were (an experimental group with a pretest, M=23.36, SD=2.987; an experimental group with no pretest, M=24.85, SD=2.617); and the comparison groups scored (comparison group with a pretest, M=16.34, SD=3.608; a comparison group with no pretest, M=16.16, SD=3.148). From Table 7, mean scores of both experimental groups were greater than mean scores of comparison groups. Both experimental and comparison groups had a wide range of scores among their subjects. Their distributions, however, were comparable among groups. A one-way ANOVA was used to determine differences between the mean scores, and as presented in Table 8.

As it has been presented in Table 8, the mean score differences between and within groups had shown statistically significant difference, $F (3.161)=88.568$, $p<0.05$. Tukey's post hoc test was performed to discover where statistically significant differences occurred in post-test mean scores. The test results in Table 9 show a significant difference in post-test performance among pre-tested experimental groups and pre-tested comparison groups, with $p<0.05$. A statistically significant difference was also found between the pre-tested comparison group, the pre-tested experimental group, and the non-pre-tested experimental group. There was no statistically significant difference in post-test mean scores between the two experimental groups, pre-tested and non-pre-tested, which may be attributed to the intervention's similar effects on their learning. In terms of post-test mean scores, no significant difference exists among the two-comparison group, the pre-tested comparison group and the non-pre-tested comparison group. There were not statistically significant differences between the groups that took the pre-test and those that did not as seen in Table 9.

Any disparities discovered can therefore be attributable to the instructional strategy. The posttest means for the comparison (pre-tested and non-pre-tested) and experimental (pre-tested and non-pre-tested) groups were then combined. The combined composition and experimental groups were created and utilized in the ensuing post-test analysis of overall performance. An independent sample t-test was used to compare the mean scores of the combined experimental and control groups as revealed in Table 10.

As seen from Table 10, achievement mean scores of the combined groups differed. The combined experimental group had a greater achievement test score (M=24.14, SD=2.87) than the combined comparison group's achievement test score (M=16.25, SD=3.36). Table 11 depicts an independent sample t-test was employed to see if the observed differences in mean achievement test scores were statistically significant. There were statistically significant variations in post-test achievement mean scores between the two groups ($t (163)=16.05$, $p<0.05$). Learners in the experimental group had superior achievement test score over the comparison group.

Table 7. Descriptive statistics of post-achievement test

	N	Mean	SD	SE	95% confidence interval for mean Lower bound	Upper bound
Experimental group with pretest	36	23.36	2.98	.498	22.35	24.37
Comparison group with pretest	44	16.34	3.60	.544	15.24	17.44
Experimental group with no pretest	40	24.85	2.61	.414	24.01	25.69
Comparison group with no pretest	45	16.16	3.14	.469	15.21	17.10
Total	165	19.88	5.04	.393	19.11	20.66

Table 8. Results of ANOVA for the post-achievement test

	Sum of squares	df	Mean square	F	Sig.
Between groups	2599.609	3	866.536	88.568	.000
Within groups	1575.203	161	9.784		
Total	4174.812	164			

Table 9. Post hoc analysis of differences in the post-achievement test scores

(I) Group of respondents	(J) Group of respondents	Mean difference (I-J)	SE	Sig.
Experimental group with pretest	Comparison group with pretest	7.020*	.703	.000
	Experimental group with no pretest	-1.489	.719	.167
	Comparison group with no pretest	7.206*	.699	.000
Comparison group with pretest	Experimental group with pretest	-7.020*	.703	.000
	Experimental group with no pretest	-8.509*	.683	.000
	Comparison group with no pretest	.185	.663	.992
Experimental group with no pretest	Experimental group with pretest	1.489	.719	.167
	Comparison group with pretest	8.509*	.683	.000
	Comparison group with no pretest	8.694*	.680	.000
Comparison group with no pretest	Experimental group with pretest	-7.206*	.699	.000
	Comparison group with pretest	-.185	.663	.992
	Experimental group with no pretest	-8.694*	.680	.000

*. The mean difference is significant at the 0.05 level

Table 10. Post-test mean scores of combined experimental and comparison groups

	Comparison and experimental groups	N	Mean	SD	SE
Achievement post-test	Experimental group	76	24.14	2.878	.330
	Comparison group	89	16.25	3.365	.357

Table 11. Independent sampled t-test for the combined experimental and comparison groups

		Levene's test for equality of variances		t-test for equality of means				
		F	Sig.	T	df	Sig. (2-tailed)	MD	SED
Achievement posttest	Equal variances assumed	2.13	.14	16.05	163	.000	7.898	.492
	Equal variances not assumed			16.25	162.9	.000	7.898	.486

3.3.2. Analysis of the effect of gender on students' achievement during technology-based instruction

An independent sample t-test was done using the post-achievement score of the treatment group to see whether the effect of technology-based training is gender-dependent. Table 12 emphasizes the results of the independent t-test on achievement scores of the combined treatment groups. The table summarizes the findings of the independent samples t-test comparing combined (pre-tested and non-pre-tested) experimental groups of males and females on post-achievement test scores. Results of an independent t-test between combined boys and girls on the post-achievement test results are highlighted in Table 12.

Table 12. Post-test mean scores of male and female (combined experimental groups)

	Gender	N	Mean	SD	S E
Post-achievement test score	Male	39	23.97	2.75	.441
	Female	37	24.32	3.02	.497

As seen in Table 12, the mean achievement test score of males (M=23.97, SD=2.75) was lower than females (M=24.32, SD=3.02). An independent sampled t-test was conducted to determine whether the difference was statistically significant. As shown in Table 13, the independent sampled t-test scores at $t(74)=-0.527$; $p>0.05$, which is higher than 0.05 ($p>0.05$). This means there was no statistically significant academic achievement differences between male and female students during technology-based instruction. As a result, when technology is used in the classroom, both boys and girls can benefit from technology integrated instruction.

Table 13. Independent sample t-test between males and females (combined experimental) on the post-test results

		Levene's test for equality of variances		t-test for equality of means				
		F	Sig.	t	df	Sig. (2-tailed)	MD	SED
Post-achievement test score	Equal variances assumed	.266	.608	-.527	74	.600	-.349	.663
	Equal variances not assumed			-.526	72.4	.601	-.349	.665

3.3.3. Analysis of the effects of intervention on students' retention of concepts

To evaluate the effect of the intervention on students' retention capacity, learners in both experimental and control groups were subjected to take a post-test after four weeks delay of the intervention/instruction. The one-way ANOVA test results of the groups' performance are shown in Table 14. According to the table, there was a mean retention test score difference between the experimental and comparison groups. In general, experimental groups retained more than comparison groups. The distributions, however, were comparable among groups. A one-way ANOVA was performed to determine whether there is statistical significance in retention mean score differences and they are shown in Table 15.

Table 14. Descriptives statistics of retention test scores

	N	Mean	SD	SE	95% confidence interval for mean Lower bound	Upper bound
Experimental group with pretest	36	21.89	2.605	.434	21.01	22.77
Comparison group with pretest	44	17.41	3.990	.602	16.20	18.62
Experimental group with no pretest	40	18.30	2.441	.386	17.52	19.08
Comparison group with no pretest	45	17.38	4.086	.609	16.15	18.61
Total	165	18.59	3.835	.299	18.00	19.18

Table 15. Results of ANOVA for the retention test scores

	Sum of squares	df	Mean square	F	Sig.
Between groups	523.876	3	174.625	14.749	.000
Within groups	1906.270	161	11.840		
Total	2430.145	164			

As demonstrated in Table 15, the retention mean scores between and within groups are statistically significant different, F (3.161)=14.75, p<0.05. Tukey's post hoc test was employed to examine between which group the difference existed in retention test mean scores. The results of the analysis are shown in Table 16. The test results show that pre-tested experimental groups have a statistically significant difference in retention capacity compared to the other groups (non-pre-tested experimental groups, pre-tested comparison groups, and non-pre-tested comparison groups) with p<0.05. As seen from Table 16, the test analysis of between the two comparison groups, the non-pre-tested experimental group and the pre-tested comparison groups, and the non-pre-tested experimental group and the non-pre-tested comparison groups, had a p-value greater than 0.05 (p>0.05). Hence, there was no statistically significant difference result was found between the two comparison groups in retention capacity test mean scores (p>0.05).

Table 16. Post hoc test analysis of differences in the retention test scores between experimental and comparison groups

(I) Group of respondents	(J) Group of respondents	Mean difference (I-J)	SE	Sig.	95% C I LB	U B
Experimental group with pretest	Comparison group with pretest	4.480*	.773	.000	2.47	6.49
	Experimental group with no pretest	3.639*	.791	.000	1.59	5.69
	Comparison group with no pretest	4.511*	.769	.000	2.51	6.51
Comparison group with pretest	Experimental group with pretest	-4.480*	.773	.000	-6.49	-2.47
	Experimental group with no pretest	-.841	.752	.679	-2.79	1.11
	Comparison group with no pretest	.031	.730	1.00	-1.86	1.93
Experimental group with no pretest	Experimental group with pretest	-3.639*	.791	.000	-5.69	-1.59
	Comparison group with pretest	.841	.752	.679	-1.11	2.79
	Comparison group with no pretest	.872	.748	.649	-1.07	2.81
Comparison group with no pretest	Experimental group with pretest	-4.511*	.769	.000	-6.51	-2.51
	Comparison group with pretest	-.031	.730	1.00	-1.93	1.86
	Experimental group with no pretest	-.872	.748	.649	-2.81	1.07

*. The mean difference is significant at the 0.05 level

This comparable retention mean scores in comparison groups can also be connected to a similar effect of teaching and learning that does not employ technology-based instruction. The pre-test did not affect the mean retention test scores, according to post hoc analysis, as there were no statistically significant differences between the comparison groups that took the pre-test and those that did not. As a result, any differences observed can be traced to the method of instruction. The retention means of the experimental (pre-tested and non-pre-tested) and comparison (pre-tested and non-pre-tested) groups were combined. The combined experimental and comparison groups were formed, and they were used in the later assessment of total idea retention ability. The mean scores of the combined experimental and combined comparison groups were subjected to an independent sample t-test as seen in Table 17.

Table 17. Retention test mean scores of combined experimental and comparison groups

	Comparison and experimental groups	N	Mean	SD	SEM
Retention test scores	Experimental group	76	21.66	3.177	.364
	Comparison group	89	17.39	4.016	.426

Table 17 depicts that retention test scores of the combined groups differed. Retention mean score in the combined experimental group was superior (M=21.66, SD=3.177) to the combined comparison group (M=17.39, SD=4.016). The independent sample t-test was used to determine if the observed mean scores differences were statistically significant. Meanwhile, Table 18 presents that there was a statistically significant difference between retention test mean score of experimental and comparison groups (t (162.09)=7.610, p<0.05). The experimental group's mean retention capacity test score was statistically higher than that of the comparison groups.

Table 18. Independent sampled t-test for the combined experimental and comparison groups

		Levene's test for equality of variances		t-test for equality of means				
		F	Sig.	t	df	Sig. (2-tailed)	MD	SED
Retention test scores	Equal variances assumed	6.432	.012	7.47	163	.000	4.265	0.571
	Equal variances not assumed			7.61	162.09	.000	4.265	0.560

3.3.4. Analysis of the effect of gender on students' retention during technology-based instruction

An independent sample t-test was performed on the retention test scores of the experimental groups to see if the effect of technology-based instruction is gender-dependent or not. The results of the independent t-test on the retention test scores of the combined (pre-tested and non-pre-tested) experimental groups are emphasized in Table 19. Finally, an independent t-test was performed using retention test scores to show whether there was a statistically significant difference in retention capacity between males and females in the combined experimental groups.

Table 19. Post-retention test mean scores of combined experimental male and female students

	Gender	N	Mean	SD	SEM
Retention test scores	Male	39	20.05	3.128	.501
	Female	37	19.89	3.187	.524

From Table 19, male retention test scores (M=20.05, SD=3.128) were higher than female retention test scores (M=19.89, SD=3.187). An independent sample t-test was conducted to determine whether or not the observed differences in mean scores were statistically significant. Also, there was no statistically significant difference in retention capacity between males and females at t (74)=0.220; p>0.05 as seen in Table 20. As a result, when technology is used in the classroom, both boys and girls benefited in the same way.

Table 20. Independent samples t-test of retention test scores between experimental male and females

		Levene's test for equality of variances		t-test for equality of means				
		F	Sig.	t	df	Sig. (2-tailed)	MD	SED
Retention test scores	Equal variances assumed	.082	.775	.220	74	.826	.159	.724
	Equal variances not assumed			.220	73.62	.827	.159	.725

4. DISCUSSIONS

The impact of technology-integrated chemistry instruction on academic achievement and retention capacity of secondary school students was examined in this study using the Solomon's four group design. According to the current study, the experimental groups performed better in terms of chemistry achievement and recall ability than the comparison groups. The dual coding theory developed by Sadoski and Paivio [48] is cited by those who favor employing modern technology to enhance learning. They contend that while traditional education approaches only use vocal presentation of material, technology-integrated apps mix text, image, audio, and video [49], [50]. According to the dual coding theory, a person can recall either words or images separately or both at once. Separately coded stimuli are less likely to be forgotten than mixed stimuli. Multi-sensory (verbal and visual) processing is developed through technology-integrated educational methods. The findingsof this study is in line with those of other studies (e.g., Poripo [51], Tolani-Brown *et al*. [52], Mahmood and Mirza [53]), which showed that students who had been taught using simulations achieved statistically better mean scores and retained more information than those who were taught traditionally. However, this study's results went against those of other studies (such as those by Wegner *et al*. [54], Owusu *et al.* [55], Imhanlahimi and Imhanlahimi [56], Hannel and Cuevas [57]), showing that technology-integrated teaching has no appreciable impact on students' academic performance.

Similar study results are reported by different researchers and professionals in education (e.g., Tolani-Brown *et al.* [52], Demissie *et al.* [1]) and Ishaq *et al*. [58] showed the beneficial effects of technology-assisted instruction on students' achievements. According to post-test findings, students in the treatment group comprehended chemical bonding concepts better than students in the comparison group, according to a study by Frailich *et al*. [10]. This test was designed to assess your ability to comprehend the links between microscopic structure and macroscopic features, as well as the different chemical bonds (metallic, ionic, and covalent). According to the results, chemical bonding-related online activities based on visualization tools (such animations, computer models, and applets) aid in better understanding abstract concepts. They found that students understood chemical concepts more effectively than non-users when they were exposed to animations and visualizations in their chemistry classes.

In a study by Mihindo *et al.* [16] and Oladejo *et al.* [59], the effect of gender on computer-based simulation on problem-solving abilities in chemistry was examined. Their research found no appreciable disparities in the performance of stoichiometric chemistry problems by male and female high school students. In opposition to this assertion, Nisrina *et al.* [60] and Chinwe and Onyebuchi [61] found that during technology-based instruction, male students performed better on average than female students. The t-test result in the current study similarly showed that no significant difference between the achievements mean scores of male and female students. This exemplifies how gender is taken into account in simulation. As a result, gender stereotypes in science education are diminished when chemistry fundamentals are taught through simulation.

Additionally, the current study's findings demonstrated that no statistically significant difference between male and female students' achievement and retention during technology-based training. These findings are in line with those of earlier studies (such as those by Tukura *et al.* [20], Ani *et al.* [24]), which showed that gender equality is a benefit of technology-based teaching. Other studies that refuted this conclusion were those by Nisrina *et al.* [60], Zubaidah *et al.* [62], Anagbogu and Ezeliora [18], and others that showed male students did better than female students.

Studies like Wegner *et al.* [54], Owusu *et al.* [55], and Hannel and Cuevas [57], which disputed the findings of the current study, were carried out utilizing a quasi-experimental research approach. The underlying flaw with a quasi-experimental design is that confounding bias cannot be completely eliminated, which leaves room for doubt in the study's conclusions. This problem is commonly brought up to challenge the results of quasi-experimental studies. While the current study used a Solomon's four-group quasi-experimental design, this research approach eliminates the impact of confounding (extraneous) variables by blocking pre-test sensitization. As a result, the findings of the current study are more likely to be reliable than those of earlier studies.

Similar to this, Imhanlahimi and Imhanlahimi's study [56] used a real experimental research design with only 60 students who were randomly selected and divided into two groups. While 165 students were sampled for the current study, which used a Solomon's group quasi-experimental design, the higher the sample size, the more accurate the average values of the results. The results of the current study are more reliable since larger sample numbers aid in identifying data outliers and offer lower margins of error.

5. CONCLUSION

When compared to the comparison groups, the experimental group's learners' mean test scores on the post-tests given immediately following technology-integrated instruction and four weeks later are significantly higher. These results therefore support the claim that technology-integrated instruction may improve academic achievement and concept memory in chemical bonding. According to the findings of the study, establishing students' idea maps increases students' achievement and promotes the retention of chemical bonding concepts. Visual representations of relationships can help students extract meaning from information by manipulating it and explaining difficult-to-understand text and abstract concepts. Concept maps can assist students in understanding the main ideas and how they are related, as well as revealing misconceptions about understanding. Even if there is disagreement that males and females benefit differently from technology-integrated instruction, this study revealed that both males and female students achieve and retain knowledge similarly during technology-integrated instruction. Both male and female students benefitted equally (no gender bias) from technology-based instructions.

ACKNOWLEDGEMENTS

We would like to extend our appreciation to the 11th-grade Damot Secondary School students and teachers who were involved in this research.

REFERENCES

[1] T. Demissie, C. E. Ochonogor, and T. Engida, "Effects of technology driven pedagogy applications on the comprehension of complex and abstract concepts of chemical equilibrium," *African Journal of Chemical Education*, vol. 3, no. 2, pp. 57–75, 2013.

[2] R. Lavi, G. Shwartz, and Y. J. Dori, "Metacognition in chemistry education: A literature review," *Israel Journal of Chemistry*, vol. 59, no. 6–7, pp. 583–597, Jun. 2019, doi: 10.1002/ijch.201800087.

[3] J. Wang, D. Guo, and M. Jou, "A study on the effects of model-based inquiry pedagogy on students' inquiry skills in a virtual physics lab," *Computers in Human Behavior*, vol. 49, pp. 658–669, 2015, doi: 10.1016/j.chb.2015.01.043.

[4] S. Ghavifekr and W. A. W. Rosdy, "Teaching and learning with technology: effectiveness of ICT integration in schools," *International Journal of Research in Education and Science*, vol. 1, no. 2, p. 175, Mar. 2015, doi: 10.21890/ijres.23596.

[5] F. K. Sarfo and K. Ansong-Gyimah, "Ghanaian senior high school students' access to and experiences in the use of information and communication technology," *Education in a technological world: communicating current and emerging research and technological efforts*, pp. 216–223, 2011.

[6] L. Cardellini, "Chemistry: Why the subject is difficult?," *Educación Química*, vol. 23, pp. 305–310, May 2012, doi: 10.1016/S0187-893X(17)30158-1.

[7] C. Nwagbo, "Effects of two teaching methods on the achievement in and attitude to biology of students of different levels of scientific literacy," *International Journal of Educational Research*, vol. 45, no. 3, pp. 216–229, Jan. 2006, doi: 10.1016/j.ijer.2006.11.004.

[8] A. Sadik, "Digital storytelling: a meaningful technology-integrated approach for engaged student learning," *Educational Technology Research and Development*, vol. 56, no. 4, pp. 487–506, Aug. 2008, doi: 10.1007/s11423-008-9091-8.

[9] M. Oubibi *et al.*, "Advances in research on technological, pedagogical, didactical, and social competencies of preservice TCFL teachers," *Sustainability*, vol. 14, no. 4, Feb. 2022, doi: 10.3390/su14042045.

[10] M. Frailich, M. Kesner, and A. Hofstein, "The influence of web-based chemistry learning on students' perceptions, attitudes, and achievements," *Research in Science & Technological Education*, vol. 25, no. 2, pp. 179–197, Jul. 2007, doi: 10.1080/02635140701250659.

[11] M. J. Sanger and S. M. Badger, "Using computer-based visualization strategies to improve students' understanding of molecular polarity and miscibility," *Journal of Chemical Education*, vol. 78, no. 10, p. 1412, 2001, doi: 10.1021/ed078p1412.

[12] H. Özmen, "The influence of computer-assisted instruction on students' conceptual understanding of chemical bonding and attitude toward chemistry: A case for Turkey," *Computers & Education*, vol. 51, no. 1, pp. 423–438, Aug. 2008, doi: 10.1016/j.compedu.2007.06.002.

[13] R. L. Palculict, "The effects of computer-assisted instruction on student learning of fractions in middle school mathematics," Doctoral Dissertations and Projects, Liberty University, 2022.

[14] C. Nja, K. J. Idiege, and O. J. Joy, "Effect of the use of ICT on chemistry performance," *Education for Today*, vol. 13, no. 1, pp. 148–153, 2017.

[15] E. Benson, "Effect of computer-aided instruction on junior secondary students'achievement and retention in basic science," Benue State University, 2016.

[16] W. J. Mihindo, S. Wachanga, and Z. Anditi, "Effects of computer-based simulations teaching approach on students' achievement in the learning of chemistry among secondary school students in Nakuru Sub County, Kenya," *Journal of Education and Practice*, vol. 8, no. 5, pp. 65–75, 2017.

[17] O. E. Nkemakolam, O. F. Chinelo, and M. C. Jane, "Effect of computer simulations on secondary school students' academic achievement in chemistry in Anambra state," *Asian Journal of Education and Training*, vol. 4, no. 4, pp. 284–289, 2018, doi: 10.20448/journal.522.2018.44.284.289.

[18] M. A. Anagbogu and B. Ezeliora, "Prospects and problems of Nigerian women in science and technology of national development," *Ghana Journal of Development Studies*, vol. 5, no. 2, 2008, doi: 10.4314/gjds.v5i2.35072.

[19] E. Yanarates, "The effect of animated teaching on science teacher candidates' chemistry achievements and learning persistence," *Educational Policy Analysis and Strategic Research*, vol. 17, no. 1, pp. 108–126, 2022, doi: 10.29329/epasr.2022.248.6.

[20] T. C. Saidu, A. Abdulmalik, and K. Judith, "Effects of e-learning on retention and performance of basic science and technology students in Minna, Niger State," *International Journal of Research and Scientific Innovation (IJRSI)*, vol. 7, no. 11, pp. 33–38, 2020.

[21] S. P. Carter, K. Greenberg, and M. S. Walker, "The impact of computer usage on academic performance: Evidence from a randomized trial at the United States Military Academy," *Economics of Education Review*, vol. 56, pp. 118–132, 2017, doi: 10.1016/j.econedurev.2016.12.005.

[22] S. P. Carter, K. Greenberg, and M. S. Walker, "Should professors ban laptops?," *Education Next*, vol. 17, no. 4, pp. 68–74, 2017.

[23] M. O. Yusuf and A. O. Afolabi, "Effects of computer assisted instruction (CAI) on secondary school students' performance in Biology," *Turkish Online Journal of Educational Technology*, vol. 9, no. 1, pp. 62–69, 2010.

[24] I. M. Ani, C. A. Obodo, C. C. Ikwueze, and I. T. Festus, "Effect of gender on basic science students academic achievement in secondary schools in Enugu Education Zone, Enugu State, Nigeria," *UNIZIK Journal of Educational Research and Policy Studies*, vol. 2, no. 1, pp. 9–14, 2021, doi: 10.5897/unijerps2021.0004.

[25] K. A. Achuonye and C. N. Olele, "Internet using patterns of Nigerian teacher-trainees: Implications for teacher education in Nigeria," *The journal of the Science Teachers' Association of Nigeria*, vol. 44, no. 1, pp. 103–108, 2009.

[26] S. R. Pate, “Effect of digital-learning on academic achievement of secondary school students.” pp. 1–33, 2019. [Online]. Available: https://www.researchgate.net/publication/339898989_Effect_of_Digital-learning_on_Academic_achievement_of_secondary_school_students.
[27] H. Cevahir, M. Özdemir, and M. H. Baturay, “The effect of animation-based worked examples supported with augmented reality on the academic achievement, attitude and motivation of students towards learning programming,” *Participatory Educational Research*, vol. 9, no. 3, pp. 226–247, 2022, doi: 10.17275/per.22.63.9.3.
[28] A. Unal and Z. Yerlikaya, “The effect of using social networks in the inquiry-based general chemistry laboratory course,” *International Journal of Progressive Education*, vol. 17, no. 6, pp. 16–31, 2021, doi: 10.29329/ijpe.2021.382.2.
[29] S. Çınar and S. Çepni, “The impact of science teaching based on science-technology-society (STS) approach to elementary school students,” *Educational Policy Analysis and Strategic Research*, vol. 16, no. 4, pp. 198–217, 2021, doi: 10.29329/epasr.2021.383.11.
[30] A. Yilmaz and E. Yanarateş, “The effect of STEM activities developed within the scope of a science course on 7th grade students’ inquiry and innovative thinking skills: The effect of STEM activities,” *International Journal of Curriculum and Instruction*, vol. 14, no. 1, pp. 274–303, 2022.
[31] A. Iyamuremye, J. Mukiza, E. Nsabayezu, F. Ukobizaba, and K. Ndihokubwayo, “Web-based discussions in teaching and learning: Secondary school teachers’ and students’ perception and potentiality to enhance students’ performance in organic chemistry,” *Education and information technologies*, vol. 27, no. 2, pp. 2695–2715, 2022, doi: 10.1007/s10639-021-10725-7.
[32] D. Ardac and S. Akaygun, “Effectiveness of multimedia-based instruction that emphasizes molecular representations on students’ understanding of chemical change,” *Journal of Research in Science Teaching*, vol. 41, no. 4, pp. 317–337, 2004, doi: 10.1002/tea.20005.
[33] M. M. Ratamun and K. Osman, “The effectiveness comparison of virtual laboratory and physical laboratory in nurturing students’ attitude towards chemistry,” *Creative Education*, vol. 09, no. 09, pp. 1411–1425, 2018, doi: 10.4236/ce.2018.99105.
[34] M. E. Yesilyurt, R. Basturk, F. Yesilyurt, and I. Kara, “The effect of technological devices on student’s academic success: evidence from Denizli,” *Journal of Internet Applications and Management*, vol. 5, no. 1, pp. 39–47, 2014, doi: 10.5505/iuyd.2014.83007.
[35] F. Jabeen and M. T. Afzal, “Effect of simulated chemistry practicals on students’ performance at secondary school level,” *Journal of Education and Educational Development*, vol. 7, no. 1, p. 119, 2020, doi: 10.22555/joeed.v7i1.2600.
[36] M. D. Miller, G. Castillo, N. Medoff, and A. Hardy, “Immersive VR for organic chemistry: impacts on performance and grades for first-generation and continuing-generation university students,” *Innovative Higher Education*, vol. 46, no. 5, pp. 565–589, 2021, doi: 10.1007/s10755-021-09551-z.
[37] I. Hussain, Q. Suleman, M. Naseer-ud-Din, and F. Shafique, “Effects of information and communication technology (ICT) on students’ academic achievement and retention in chemistry at secondary level,” *Journal of Education and Educational Development*, vol. 4, no. 1, p. 73, 2017, doi: 10.22555/joeed.v4i1.1058.
[38] C. Andrade, “The limitations of quasi-experimental studies, and methods for data analysis when a quasi-experimental research design is unavoidable,” *Indian journal of psychological medicine*, vol. 43, no. 5, pp. 451–452, Sep. 2021, doi: 10.1177/02537176211034707.
[39] N. N. Mai, Y. Takahashi, and M. M. Oo, “Testing the effectiveness of transfer interventions using Solomon four-group designs,” *Education Sciences*, vol. 10, no. 4, p. 92, 2020, doi: 10.3390/educsci10040092.
[40] T. E. Dawson, “A primer on experimental and quasi-experimental design,” in *Annual Meeting of The Southwest Educational Research Association*, 1997. [Online]. Available: https://files.eric.ed.gov/fulltext/ED406440.pdf.
[41] T. W. McGahee and M. S. Tingen, “The use of the solomon four-group design in nursing research,” *Southern Online Journal of Nursing Research*, vol. 9, no. 1, 2009.
[42] R. L. Solomon, “An extension of control group design,” *Psychological Bulletin*, vol. 46, no. 2, pp. 137–150, 1949, doi: 10.1037/h0062958.
[43] S. R. Aschen, “Assertion training therapy in psychiatric milieus,” *Archives of Psychiatric Nursing*, vol. 11, no. 1, pp. 46–51, 1997, doi: 10.1016/s0883-9417(97)80049-0.
[44] S. E. Dare, W. W. van Dijk, E. van Dijk, L. F. van Dillen, M. Gallucci, and O. Simonse, “The effect of financial education on pupils’ financial knowledge and skills: Evidence from a Solomon four-group design,” *The Journal of Educational Research*, vol. 113, no. 2, pp. 93–107, 2020, doi: 10.1080/00220671.2020.1733453.
[45] D. S. Whitman, D. V. Rooy, C. (Vish) Viswesvaran, and A. Alonso, “The susceptibility of a mixed model measure of emotional intelligence to faking: a Solomon four-group design,” *Psychology Science Quarterly*, vol. 50, no. 1, pp. 44–63, 2008.
[46] G. Ursachi, I. A. Horodnic, and A. Zait, “How reliable are measurement scales? External factors with indirect influence on reliability estimators,” *Procedia Economics and Finance*, vol. 20, pp. 679–686, 2015, doi: 10.1016/s2212-5671(15)00123-9.
[47] D. M. Mertens, “Mixed methods and wicked problems,” *Journal of Mixed Methods Research*, vol. 9, no. 1, pp. 3–6, Jan. 2015, doi: 10.1177/1558689814562944.
[48] M. Sadoski, M. Sadoski, A. Paivio, and A. Paivio, *Imagery and text: A dual coding theory of reading and writing*. New York: Routledge, 2013, doi: 10.4324/9781410605276.
[49] C. Nicolaou, M. Matsiola, and G. Kalliris, “Technology-enhanced learning and teaching methodologies through audiovisual media,” *Education Sciences*, vol. 9, no. 3, p. 196, 2019, doi: 10.3390/educsci9030196.
[50] C. Tarchi, S. Zaccoletti, and L. Mason, “Learning from text, video, or subtitles: A comparative analysis,” *Computers & Education*, vol. 160, p. 104034, 2021, doi: 10.1016/j.compedu.2020.104034.
[51] J. Poripo, “Effects of simulation on male and female students’ achievement in chemistry,” *Bayelsa State, Nigeria. International Journal of Education and Practice*, vol. 2, no. 1, pp. 44–51, 2008.
[52] N. Tolani-Brown, M. McCormac, and R. Zimmermann, “An analysis of the research and impact of ICT in education in developing country contexts,” *ICTs and Sustainable Solutions for the Digital Divide*. IGI Global, pp. 218–242, 2011, doi: 10.4018/978-1-61520-799-2.ch011.
[53] M. K. Mahmood and M. S. Mirza, “Effectiveness of computer-assisted instruction in Urdu language for secondary school students’ achievement in science,” *Language in India*, vol. 12, no. 2, pp. 266–283, 2012.
[54] S. B. Wegner, K. C. Holloway, and E. M. Garton, “The effects of internet-based instruction on student learning,” *Online Learning*, vol. 3, no. 2, 2019, doi: 10.24059/olj.v3i2.1920.
[55] K. A. Owusu, K. A. Monney, J. Y. Appiah, and E. M. Wilmot, “Effects of computer-assisted instruction on performance of senior high school biology students in Ghana,” *Computers & Education*, vol. 55, no. 2, pp. 904–910, 2010, doi: 10.1016/j.compedu.2010.04.001.
[56] E. O. Imhanlahimi and R. E. Imhanlahimi, “An evaluation of the effectiveness of computer assisted learning strategy and

expository method of teaching biology: a case study of lumen Christi international high school, Uromi, Nigeria," *Journal of Social Sciences*, vol. 16, no. 3, pp. 215–220, 2008, doi: 10.1080/09718923.2008.11892621.
[57] S. L. Hannel and J. Cuevas, "A study on science achievement and motivation using computer-based simulations compared to traditional hands-on manipulation," *Georgia Educational Researcher*, vol. 15, no. 1, 2018, doi: 10.20429/ger.2018.15103.
[58] K. Ishaq, N. A. Mat Zin, F. Rosdi, A. Abid, and M. Ijaz, "The impact of ICT on students' academic performance in public private sector universities of Pakistan," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 3, pp. 1117–1121, 2020, doi: 10.35940/ijitee.c8093.019320.
[59] A. I. Oladejo, N. C. Nwaboku, P. A. Okebukola, and I. A. Ademola, "Gender difference in students' performance in chemistry – can computer simulation bridge the gap?," *Research in Science & Technological Education*, vol. 41, no. 3, pp. 1031–1050, 2021, doi: 10.1080/02635143.2021.1981280.
[60] M. Gunawan, N. Nisrina, N. M. Suranti, and R. Ekasari, "The effect of virtual labs toward students' understanding of physics based on gender," *Proceedings of the 1st International Conference on Education Innovation (ICEI 2017)*. Atlantis Press, 2018, doi: 10.2991/icei-17.2018.34.
[61] O. Chinwe and O. E. Onyebuchi, "Influence of gender, school location and the use of play-simulation on the school achievemnet in chemistry," *Journal of Research in National Development*, vol. 9, no. 1, 2011.
[62] S. Zubaidah, N. M. Fuad, S. Mahanal, and E. Suarsini, "Improving creative thinking skills of students through differentiated science inquiry integrated with mind map," *Journal of Turkish Science Education*, vol. 14, no. 4, pp. 77–91, 2017.

BIOGRAPHIES OF AUTHORS

Derejaw Yesgat is Ph.D. candidate in chemistry education at Bahir Dar University. He did his first degree in chemistry education from Addis Ababa University and second degree in organic chemistry from Bahir Dar University. He has been working as a teacher educator for more than a decade. His researchareas are technology-based chemistry teaching and technology-assisted inquiry-based chemistry teaching. He can be contacted at email: yesgatderejaw@gmail.com.

Solomon Melesse is professor in curriculum and instruction at Bahir Dar University. He has been teacher educator for the last 21 years. He published several articles related to teacher education and professional development. His main areas of research include teacher education and professional development, instructional design, curriculum studies, curriculum development, and curriculum implementation. He can be contacted at email: btlhmslmn1997@gmail.com.

Dereje Andargie earned B.E.D in Chemistry Haramaya University, MA in curriculum and instructiojn in Addis Ababa University, Ph.D. in MSTE (chemistry education) from University of South Africa. He currently works in supervising Ph.D. and M.Ed. students in chemistry in Addis Ababa and Bahidar Universities; as an assistant professor in science education (chemistry education) and a vice president for academic affairs. He can be contacted at email: derejefamaye@gmail.com.

Belete Bedemo Beyene is associate professor and a full-time staff member in the department of chemistry, Bahir Dar University, Bahir Dar Ethiopia. He did bachelor of education in chemistry, M.Sc. and Ph.D. in the field of organic chemistry. His research interest basically revolves around energy and catalysis application of chemistry, natural product chemistry and chemistry education fields. He can be contacted at email: beleteb2002@gmail.com.