EUROPEAN JOURNAL OF STEM EDUCATION

European Journal of STEM Education,
2025, 10(1), 03
ISSN: 2468-4368

Effect of Virtual Experiments Compared to Physical Experiments on Students' Conceptual Understanding of Chemical Kinetics Concepts

Tadesse Hunegnaw [1]*, Tesfaye Demisse Hailegebreal [2], Dawit Asrat Getahun [1],
Minaleshewa Atlabachew [1]

[1] *Bahir Dar University*, ETHIOPIA
[2] *Debre Berhan University,* ETHIOPIA

***Corresponding Author:** tadie2912@gmail.com

Citation: Hunegnaw, T., Hailegebreal, T. D., Getahun, D. A. and Atlabachew, M. (2025). Effect of Virtual Experiments Compared to Physical Experiments on Students' Conceptual Understanding of Chemical Kinetics Concepts. *European Journal of STEM Education*, 10(1), 03. https://doi.org/10.20897/ejsteme/16261

Published: April 9, 2025

ABSTRACT

Even though physical experiments are mandatory in science education, there has been an increasing demand for virtual experiments. This research aimed to investigate the impact of virtual experiments compared to physical experiments on students' conceptual understanding of chemical kinetics concepts. To achieve this, a two-phase explanatory sequential mixed method approach was employed to collect and analyze quantitative and qualitative data. One group of students conducted virtual experiments (n = 25), while another group of students practiced the same topic with physical experiments (n = 26). The students' concept maps were used as an assessment tool before and after the intervention. Concept maps are considered a more effective assessment tool compared to traditional tests as they better evaluate students' conceptual understanding. The independent sample t-test analysis indicates no significant difference between the two groups in the post-concept maps, which suggests that the two experiments are similar in enhancing conceptual understanding. The qualitative analyses of concept maps indicate students preferred using the macroscopic levels to represent their conceptual knowledge of chemical kinetics. Future research may focus on an effective combination of virtual and physical experiments to link sub-microscopic levels to the other two levels (macroscopic and symbolic) of conceptual understanding.

Keywords: virtual experiments, physical experiments, concept maps, conceptual understanding, chemical kinetics

INTRODUCTION

Science education is an important aspect of schools in Ethiopia, as it is in every other country around the world. According to various studies, for instance (Belay et al., 2016; Engida, 2002; Lopez et al., 2011; Millar, 2004), it is essential to provide students with a basic understanding of science as a foundation for future endeavors. In particular, high school students must complete general chemistry, which is a prerequisite for further study in STEM (science, technology, engineering, and mathematics) fields in higher education. This is important for the socio-economic development of a nation, as quality education in STEM subjects is crucial for the technological advancement of a country (Lopez et al., 2011).

Practical or laboratory work plays an essential role in science education. Laboratory activities are an important component of science learning that connects theoretical lectures with practical processes (Domínguez et al., 2018). Some of the roles of laboratory activities include supporting conceptual understanding, introducing inquiry skills, synergizing minds-on and hands-on activities, etc. Laboratory creates an opportunity for students to observe, explain, discover, and conclude in a given domain which can develop scientific thinking (Millar, 2004; Nakhleh et al., 2002). Laboratory activities also help students develop general skills such as problem-solving, teamwork, reporting, presentation, time management, etc. (Seery et al., 2019). To sum up, laboratory activity is not an 'optional extra' in learning science rather it is an essential component (Millar, 2004).

In Ethiopian high schools, the need for science to be supported with laboratory activities has been frequently emphasized within curriculum frameworks and research recommendations to promote realistic learning of science (Bekalo and Welford, 2000; Belay et al., 2016; MoE, 2009). However, there is a mismatch between what is intended to be done and the actual practice in secondary schools regarding practical activities. The studies reported that a tenuous mismatch between objectives, assessment, and actual instructional practices is revealed, resulting in students receiving insufficient practical activities. Hunegnaw and Melesse (2023) also indicated that the laboratory instructions in students' textbooks should be revisited to meet the intended objectives.

Studies in Ethiopian high schools also showed that scarcity of laboratory materials (chemicals, apparatus, and reagents) due to inadequate budget is the most common reason for the limited practice of laboratory activities in science education (Zengele and Alemayehu, 2016). In turn, Ayalew (2017) states that the lack of well-equipped laboratories is the leading factors that hinder the quality of science education in Ethiopia. One probable solution to improve science learning in laboratories is integrating digital technology (Belay et al., 2016). Relatively cheap or free virtual laboratory activities like online videos and simulations could be alternatives to practical activities but the question is to what extent they develop students' scientific knowledge.

Practical work with physical or virtual experiments could be encompassed in the cognitive theory of multimedia learning (CTML). According to this theory, knowledge is generated from the combination of visual elements (for instance, color changes and illustrations displayed on computer screens) and verbal information (like worksheets and measurement data) (Altmeyer et al., 2020; Mayer, 2005; Sorden, 2012). Richard E. Mayer is a well-known advocate of CTML who suggests that the human brain learns more effectively through multimedia. The material presented in this way helps learners construct knowledge when they are actively engaging with it (Irby et al., 2013). In this study, visuals and verbal information help students to understand concepts in both virtual and physical laboratory activities, which align with the CTML theory. Students follow the visuals of laboratory activities in addition to the verbal information provided by the experimental activities.

Although physical experiments (PEs) are broadly used in schools, digital technologies like virtual experiments (VEs) have become increasingly important in science education (Kapici et al., 2019). A study comparing virtual versus real laboratory activities showed similar learning outcomes indicating proper application of virtual experiments produces the intended learning outcomes (Winkelmann et al., 2020). These authors added that students get advantages in virtual labs because of the simplification, more entertainment, and minimization of distractions during the activities. A recent systematic review of virtual laboratories in chemistry education stated that VE exceeds passive teaching techniques in learning outcomes (Chan et al., 2021). The review further showed that similar learning outcomes compared to physical laboratory activities revealed that VEs can complement or replace PEs. VEs can direct students' concentration to a specific activity or event (Millar, 2004). VEs can be used for learning domains that could not be reproduced by PEs, and repeat experiments without extra cost. The overall cost of PEs is higher since it requires more space, human power, apparatus, and chemicals making VEs better in cases where financial constraint is a problem in developing countries like Ethiopia.

However, other studies reported that VEs could not be used in place of PEs due to their limited nature. Unlike virtual labs, PEs provide rich real data, which creates the opportunity for students to link real-world activities with mind-on activities (Puntambekar et al., 2021). A review study has demonstrated that VEs can yield positive learning outcomes in certain domains, but they cannot replace the hands-on laboratory experience. Better learning outcomes can be achieved when both VEs and PEs are used in combination (Sypsas and Kalles, 2018). Since PEs and VEs can support different abilities and experiences, complementing the two leads to better outcomes (Kapici et al., 2019; Puntambekar et al., 2021).

Conceptual understanding is defined as "relational knowledge about concepts and their interrelations in a specific science domain" (Flegr et al., 2023). Johnstone triangular representations describe chemical 'reality' as a blend of macro, micro, and symbolic components (Talanquer, 2011). The macro level is related to phenomena that are tangible and visible in our world whereas the sub-microscopic describes particulate theories and models of matter. The symbolic level refers to language of the discipline including symbols, equations, and mathematical formulas that encompass both mathematical and chemical signs. Therefore, to understand the concepts of chemistry, students need to connect interlinked sub-concepts and form networked knowledge, which requires different levels of representations from the triplet (Krajcik and Shin, 2023).

However, students frequently experience difficulty in connecting the triplet nature of chemistry concepts (Rahmawati et al., 2022; Yaman and Ayas, 2015). For instance, students experience difficulties in transforming the rate of reaction concepts within and across the macroscopic, sub-microscopic, and symbolic representations

(Cakmakci et al., 2006). Furthermore, the concept of chemical kinetics is difficult for students since it requires the skill of mathematics, interpreting data, and controlling variables during practical activities.

Most studies related to the comparison of virtual and physical laboratory activities seem to go in the same direction in using traditional assessment methods (such as multiple choices, true/false, and filling blank spaces) (Brinson, 2015; Hodges et al., 2018; Krüger et al., 2022; Penn and Ramnarain, 2019). Some scholars also stated in the limitation part of their study that frequently used traditional assessment techniques are insufficient to measure the understanding of students (Samon and Levy, 2021). For instance, in multiple-choice tests, students could write the correct answer by chance without understanding. On the other hand, a review of 56 research articles in virtual and traditional laboratories reported that most studies (71%) use tests and quizzes, practical exams (9%), and lab reports/written assignments (9%) as their assessment tool (Brinson, 2015). This type of evaluation does not give a deep understanding of the learning context (Chan et al., 2021; Chiu et al., 2015; Tatli and Ayas, 2013). In Ethiopia, tests and quizzes are also widely used as summative assessments in science classrooms (Ayalew, 2017). This type of test only requires the recall of ideas that cannot assess the nature of science. Thus, science learning in Ethiopia has turned to rote learning and the assessment focuses on recalling knowledge only to pass exams (Bekalo and Welford, 1999).

Scholars suggest moving away from rote learning and assessment in favor of interactive concept construction. Common assessment tools like multiple-choice questions limit student involvement and restrict higher-order thinking skills development. To foster such skills, instruction and assessment methods must be changed (Brinson, 2015; Nicoll, 2001). The application of authentic assessments like concept maps makes students think a little harder and prepare a bit better (Burry-Stock and Oxford, 1994) which is essential to develop not only conceptual understanding but also critical thinking.

A concept map (CM) is a structural representation composed of nodes and labeled lines. Proportion is part of a concept map that represents two consecutive nodes and labeled lines which is a meaningful basic unit to evaluate the accuracy of the relationship between the concepts (Kaya, 2008; Schwendimann, 2015; Stoddart et al., 2000). It is a generative form of assessment that is used to produce interconnection between scientific ideas which is an essential property of knowledge (Schwendimann, 2015). Teachers can use concept maps to identify learning difficulties of students that could be unnoticed by other traditional tests (Lopez et al., 2011; Ruiz-Primo and Shavelson, 1996). This helps teachers to revise their teaching style and their students' learning. Unlike traditional assessment methods, concept maps provide a comprehensive and unique perspective on students' understanding (Ruiz-Primo and Shavelson, 1996). Multiple-choice tests limit students' thinking and gauge a different type of outcome (i.e. rote learning) than the learning measured by concept maps (meaningful learning). Therefore, this research uses a concept map as an innovative assessment tool to evaluate the effect of virtual experiments compared to physical experiments on the conceptual understanding of students. A concept map was used as an assessment tool for both virtually and physically instructed groups.

This paper focuses on a less addressed but important topic based on the current literature; students' concept maps as an assessment tool for better investigations of how each mode of experiments supports conceptual learning in chemical kinetics laboratory activities in two settings (virtual and physical). Hence this study was proposed to answer the following research questions:

1. Are there significant differences in the conceptual understanding of students as assessed by concept maps between the groups instructed with physical experiments and virtual experiments in chemical kinetics concepts?
2. What kind of representational levels (macroscopic, sub-microscopic, and symbolic) are dominantly indicated by concept maps in virtual and physical experimental activities?

RESEARCH METHODOLOGY

Research Design

In this study, the conceptual understanding of students was analyzed using concept maps to investigate the effect of virtual experiments (VEs) compared to physical experiments (PEs). To achieve this, a quasi-experimental research design was employed. Two existing classes taught by the same teachers participated in the study. Guided inquiry instruction was used for laboratory activities to enhance students' curiosity and achieve learning objectives. This type of instruction is more effective than open inquiry and confirmatory instructions in supporting students before, during, and after the lesson (Gericke et al., 2023). To gain a better understanding of the student's conceptual understanding, both qualitative and quantitative approaches were used to analyze concept maps. This study uses three main procedures in constructing concept maps developed by Stoddart et al. (2000): the tasks given to the students (data collection), the scoring phase (coding and analysis of the data), and a verification phase (validity and reliability were checked).

To make it more detailed regarding the research design employed, a two-phase explanatory sequential mixed method approach was utilized. The initial phase involved collecting quantitative data, followed by gathering qualitative data (Cresswell, 2012). The prioritized quantitative data provided a comprehensive overview of the research, while the qualitative data contributed to a more detailed understanding of the overall picture. Specifically, students' concept maps were scored to obtain quantitative results for analysis. Afterward, small samples were selected from each group based on high, medium, and low scores for qualitative analysis.

Research Context

In Ethiopian high schools, chemistry is a fundamental part of the curriculum for students learning from grade levels 9 to 12 (MoE, 2009). Despite challenges in resources and access, the subject is designed to provide students with essential scientific knowledge and skills to prepare them for further higher education. Chemical kinetics is one of the six units in grade 11 chemistry textbooks. In Ethiopian high schools, students practice laboratory activities as part of science learning, integrating theoretical concepts (Hunegnaw and Melesse, 2023). In this respect, chemical kinetics practical activities were practiced after students covered the lecture (theoretical part). Except for the type of instruction, the intervention took place under usual conditions (time used assigned by the school, no extra equipment and apparatus purchased, grouping was based on their prior chemistry exam scores) following the schedule arranged by the laboratory technical expert who was responsible for the overall organization of the laboratory. Students in each class met for 3 hours, once a week, for 12 weeks for the orientation, training, and main intervention (laboratory activity). Due to limited facilities, students work in groups in physical and virtual experiments during guided inquiry instruction. The lab technician guides students for intended instruction. Concept maps were constructed individually despite all other activities performed in small groups. A worksheet simplified the lab report writing and practical activities.

Research Participants

The participants of this study were natural science students from a high school in a town, in North-west Ethiopia, who were taking general chemistry. A convenient sampling technique was utilized to select the school for easy access by the researcher. Two Grade 11 sections were selected comprehensively to participate in the study. The two sections were randomly assigned as experimental and control groups. The experimental group (EG) practiced virtual experiments, whereas the control group (CG) practiced physical laboratory activities. The CG and EG differed only in the practice of chemical kinetics activities, virtually or physically.

Small groups were formed within the two sections to conduct practical activities, lab report writing, and additional activities during the intervention. Some students expected to participate in the study did not complete the full tasks due to different reasons like co-curricular activities and absenteeism for unknown reasons. Out of the 75 students who had signed the free consent form to participate, 68% had completed all the required activities and were included in the data analysis. However, research showed that a study with a concept map as an assessment greater than 30 participants is still assumed to be large enough (Watson et al., 2016). The students were predominantly male (70.59%) and had an age range of 18–21.

Two teachers were involved in instructing students, both of whom have master's degrees in chemistry. One of the teachers provided a detailed introduction to the theoretical concepts underlying the laboratory activities, ensuring that students had a strong foundation upon which to build their understanding. The other teacher (lab technician) was responsible for guiding the activities virtually and physically. The two teachers play the role of initiators, guiding students through inquiry-based activities. Initially, one teacher briefly introduces the theoretical concepts, setting the stage for the practical work to follow. Meanwhile, the lab technician is available to guide students as they engage in hands-on/virtual activities. The bulk of the class time is dedicated to students honing their skills through practical exercises under the supervision and support of the teacher and lab assistant. One of the authors of this study uses a checklist during instruction to confirm that the instruction is based on the intended plan. Students formulate a question and hypothesis, practice the experiment, and reflect on their analysis and conclusions. Approximately, students use most of the time (80%) and the teachers play a guiding role.

Physical Laboratory Instruction in this Study

Real or physical experiments involve the use of concrete, natural, or artificial objects to provide opportunities for learning about phenomena in the real world (Flegr et al., 2023). Physical or real experiments were utilized in the current study to aid in the understanding of chemical kinetics concepts. The contents include measuring the speed of reaction, the effect of changes in temperature on rate, the effect of changes in concentration, the effect of changes in surface area, the effect of the nature of reactants, and the effect of the presence of a catalyst. The physical laboratory activities followed the students' textbook laboratory procedures with modifications to align with guided inquiry instructions, including asking questions, formulating hypotheses, data collection and analysis, and reflections on the practiced activities.

Physical laboratory activities during this research made use of authentic chemicals and apparatus as specified in their textbooks. The chemicals employed include HCl, $CaCO_3$, $Na_2S_2O_3$, MnO_2 (as a catalyst), Cu, Mg, and more. Similarly, the apparatus employed during these activities includes beakers, volumetric flasks, thermometers, test tubes, measuring cylinders, connecting tubes, etc. In the experiment measuring reaction rate, for instance, the reaction of HCl and $CaCO_3$ is demonstrated to generate CO_2, which is monitored through the loss of mass of the system. In this experimental activity, apparatuses such as a volumetric flask, beam balance, cotton, and stand were used.

Virtual Laboratory Instruction in this Study

Virtual experiments offer new perspectives in laboratory activities. Simulations, animations, and videos allow users to manipulate variables and observe their effects. Dynamic graphs, diagrams, macroscopic events, and microscopic events can be illustrated with virtual laboratories (Flegr et al., 2023; Yaman and Ayas, 2015).

The laboratory room in the school under study has an internet connection with desktop computers and high-definition television which is used to facilitate online virtual chemical kinetics laboratory activities. Different sources from the web were used to practice the virtual experiments. In this study, different simulations and videos that resemble variable inputs of physical experiments from free online sources were used. These include PhET Interactive Simulations (https://phet.colorado.edu/), OLabs (https://www.olabs.edu.in/), simulations from American Association of Chemistry Teachers (https://teachchemistry.org/classroom-resources/collections/aact-simulations), Royal Society of Chemistry (https://www.youtube.com/watch?v=tWS_ohF543g) were employed to perform virtual experiments. Guiding questions are also used to assist students in applying the steps of guided inquiry instructions, such as formulating a question, collecting data while using simulation, and reflecting their ideas on reflective worksheets for all steps of instruction. Students practiced the reaction coordinate using single and many collisions, the effect of temperature on the rate of reaction, and the effect of concentration on the rate of reactions. The instruction (guided inquiry) and virtual labs are similar to the experiment done on the physical laboratory group to teach the experiment that variable inputs were the same as each physical experiment. The simulation displayed visualizations such as concentration versus time graphs, rate versus time graphs, concentration versus rate graphs, etc.

Validity and Reliability of the Instrument

To ensure the reliability of the scoring rubric used in the study, a rigorous analysis was conducted before the actual data collection. To achieve this, a small sample of 10 students from grade 12 in the same school where the study was conducted was selected. These students did not participate in the actual study but were used to test the rubric. The analysis involved a careful examination of the rubric's criteria, scoring system, and overall effectiveness in measuring the intended outcomes. The results of this analysis were used to refine and improve the rubric before its use in the study. Students were presented with important points about concept maps and sample concept maps. They were then asked to individually prepare their concept maps.

While using concept maps as an assessment, the issue of scoring them is very important to ensure reliable scores (Ruiz-Primo and Shavelson, 1996). To test the reliability of the results of the pilot test, ten concept maps were evaluated by two judges. The first scorer was one of the researchers and the other was a chemistry teacher in the school, both of whom have expert-level knowledge in both concept maps and the topic of chemical kinetics concepts. Before rating the students' concept maps, they familiarized themselves with the rubric (**Table 1**) used in this study and discussed and reviewed important articles about concept map scoring. They also discussed variations in scoring the concept maps and decided to promote future success in scoring them. The concept maps and the inter-rater reliability coefficient, calculated using the intra-class correlation coefficient, were found to have a moderate score (0.708) between the researcher and chemistry teacher, which is in the acceptable range (Smeeth and Ng, 2002).

The rubrics (**Table 1** and **Table 2**) underwent a thorough content validity check, which involved consulting with two experts in the field. Experts were university professors with extensive experience in undertaking educational research. They provided valuable feedback and suggestions for improving the rubric, which was carefully incorporated to ensure that the final version was accurate, reliable, and effective in assessing the intended learning outcomes. The input from the experts played a critical role in enhancing the quality of the rubric and ensuring that it met the highest standards of validity, fairness, and rigor.

Table 1. Rubric for scoring a concept map concept map (adapted from Schwendimann, 2015; Yaman and Ayas, 2015)

Criteria	Discussion of criteria	Points
Sound proposition (SP)	If a relation between two concepts exists and the connection is scientifically valid. There is a meaningful connection between the concepts.	4
Partially sound proportion (PSP)	If a relation between two concepts exists and some part of the connection is scientifically correct.	3
Partially incorrect proposition/ misconception (MS)	If a relation between two concepts exists and they do not correspond to scientific theory.	2
Wrong proposition (WP)	If a relation between two concepts exists and the proposition is scientifically wrong or irrelevant or logically no connection.	1
Blank (B)	If there is no explanation on the link.	0

Key: Total score = 4SP + 3PSP + 2MS + WP

Data Collection Procedures

The actual data collection was performed as pre- and post-concept maps before and after completing experimental activities virtually for the experimental group and physically for the control group. The concept map task was designed to the lowest level of directness except a few concepts were supplied as initiation and guidelines for concept map creation.

Concept-mapping skills are crucial for reliable results. Limited skills lead to limited understanding (Kaya, 2008). Since students never used concept maps before, they had taken training about concept maps. Despite the shortage of time for practice, a couple of weeks with six hours were spent training students in constructing concept maps. The training session focuses on the purpose, definition of important terms, and ways of constructing concept maps. There is an emphasis on the essential components of concept maps such as concepts, descriptive linking lines, proportions, and examples. To help students understand the importance of concept map structure, examples of concept maps created by experts were displayed. Open-ended activities were used to design concept maps in the study, allowing students to explore the concepts and their connections more deeply. Using procedures developed by Watson et al. (2016), students created concept maps using the previously studied topic (chemical bonding). Practice maps were collected and feedback was given.

Before being instructed in laboratory activities, students are asked to construct their pre-concept maps guiding students to focus on chemical kinetics concepts. To help them with this task, four key concept lists were given to the students: Chemical kinetics, rate of reaction, factors affecting the rate of reaction, and temperature. After receiving these lists, the students were asked to draw concept maps by forming links and labels. Additionally, students were encouraged to use additional valid concepts while constructing the concept maps. Afterward, students engaged in practical work related to chemical kinetics concepts. They were directed to create post-concept maps independently, starting with four fundamental concepts, and later supplementing with additional concepts of their choosing. These steps were consistent across both the control and experimental groups.

Methods of Data Analysis

A mixed-method approach was used to analyze the concept maps collected both qualitatively and quantitatively to gain a better insight into students' understanding of chemical kinetics concepts. Quantitative analysis, along with qualitative investigation of concept maps, is crucial in revealing students' conceptual changes and misconceptions (Burry-Stock and Oxford, 1994). The study used pre- and post-experiment concept maps constructed by students. Quantitative analysis was performed using SPSS statistics 20 with a significance level of 0.05. The paired sample t-test compared means within a group, while the independent sample t-test compared means between groups. The concept maps were analyzed to check for differences between the groups. Post-concept maps were analyzed qualitatively to study the triplet representation of concepts in virtual chemical kinetics practical activities.

When evaluating concept maps, it is crucial to establish a fair and rational criterion. The scoring mechanism must factor in cognitive demand, content domain representation, practicality, and appropriateness (Ruiz-Primo and Shavelson, 1996). To assess propositions in concept maps with varying levels of comprehension, **Table 1** presents a rubric that has been adapted from scholarly works (Schwendimann, 2015; Yaman and Ayas, 2015). Yaman and Ayas (2015) developed a rubric to score concept maps after computer-based "predict, observe, and explain acid-base" instructions. They recommended that the rubric could be used in other domains of chemistry by teachers and researchers. As presented in **Table 1**, the current study uses a rubric that focuses on scoring the contextual information (concepts, links) in a proposition rather than the structure of concept maps (hierarchical and network). This study places greater emphasis on proportion by using a five-level scoring system ranging from

Table 2. Coding criteria for identifying the triplet nature of chemistry in students' concept maps (Gkitzia et al., 2011)

Type of representation	Description
Macro	Entities and phenomena that are tangible and visible to our world.
Sub-micro	Particulate level describing atoms, molecules, electrons, and ions.
Symbolic	Encompass both chemical and mathematical signs such as symbols, formulas, and equations.
Multiple/mixed/hybrid	Shows a chemical phenomenon simultaneously at two or three levels.

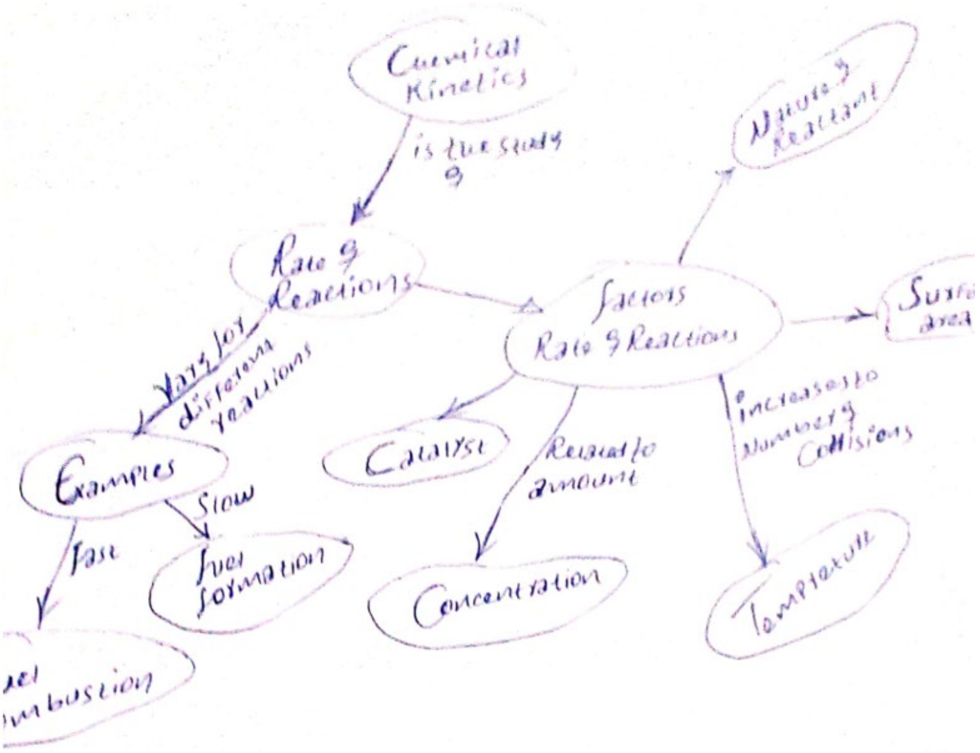

Figure 1. Sample post-concept map for a student (coded VL16)

0 to 4 points. This is due to the conceptual information is more important than the way it is presented (Nicoll, 2001).

This study categorizes proportions as sound proportion (SP), partially sound proportion (PSP), misconceptions (MS), wrong proportion (WP), and blank proportion to score 4, 3, 2, 1, and 0 points, respectively (refer to **Table 1**). The overall score is calculated using the formula 4SP + 3PSP + 2MS + WP. A sound proportion with the highest score (4) in a concept map reflects a well-understood and scientifically valid relationship between two concepts. Students' misconceptions could be revealed from the linkage of unrelated concepts. A blank (without a link) between two closely related concepts also indicates a poor conceptual understanding among students.

The qualitative analysis of concept maps is based on a small number of students from both the control and the experimental groups. To maintain validity (Jaber and BouJaoude, 2012), a representative sample with low, middle, and high concept map scores was selected. Six students' concept maps (three from the experimental and three from the control group) were selected and analyzed. **Table 2** presents a rubric (adopted from Gkitzia et al., 2011) that is used to investigate the triplet nature of chemistry for qualitative analysis. Using the rubric content analysis was used to evaluate students' concept maps depicting the three levels of the triplet nature of chemistry, namely, submicroscopic (particulate), macroscopic (observational), and symbolic (representational) levels.

Students' concept maps were evaluated with an in-depth coding process applied based on a rubric. The rubric assigns four labels (P4, P3, P2, and P1) to different proportions of marks ranging from 4 to 1, respectively. **Figure 1** and **Figure 2** present a sample of students' concept maps that have been scored based on the rubric presented in **Table 1**. **Figure 1** and **Figure 2** demonstrate samples of students' concept maps that were evaluated using the rubric in **Table 1**. The rubric allows for a comprehensive evaluation of each student's work and provides a clear understanding of their level of conceptual understanding.

Ethical Considerations

The research ethics committee of the College of Science at Bahir Dar University granted retrospective approval for this study, as this research involved human participants, after reviewing the data collection process. The

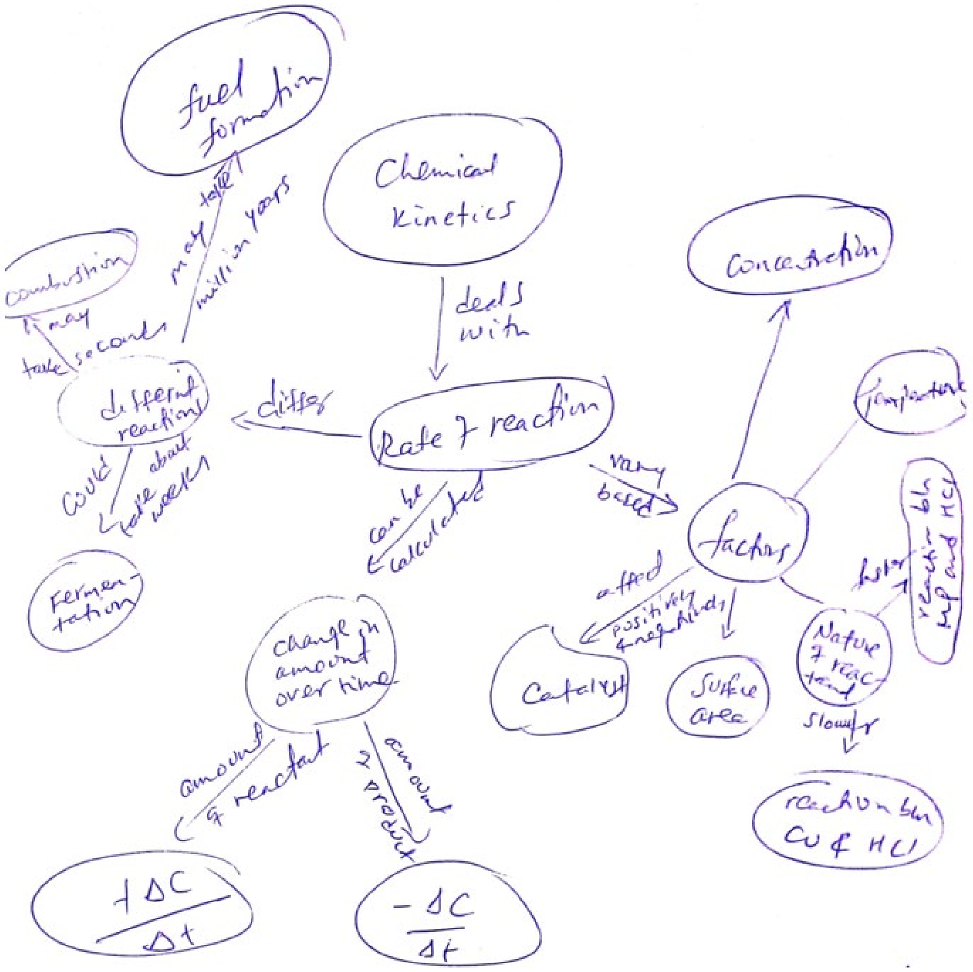

Figure 2. Post-concept map for a student (coded PL24)

Table 3. Descriptive statistics in the independent sample test for pre-concept map scores

	Lab type	N	Mean	Standard deviation	Standard error mean
PreCM	PE	26	31.00	4.501	0.883
	VE	25	33.28	5.777	1.155

participants were fully informed about the research and provided their consent through a signed form. To protect their identities, a code was used instead of their names. It is important to note that the assessment outcomes from this study were not included in their final semester grades.

RESULTS OF THE STUDY

The first objective of this study was to compare the conceptual understanding of students instructed with virtual and physical experiments. The scores of concept maps were analyzed using a t-test to determine whether there was a significant difference between the two groups. Paired sample t-test and independent sample t-test analyses were utilized for analyzing the data as it is approximately normally distributed around the mean.

A paired-sample t-test was conducted to compare the pre-concept map (mean [M] = 33.28) and post-concept map (M = 42.40) scores for the experimental group (EG). The result [t(24) = –6.003, p = 0.000] proved that there was a significant difference in the scores of the EG before and after the virtual experiment. Similarly, a paired-sample t-test was conducted to compare the pre-concept map scores (M = 31.00) and the post-concept map scores (M = 44.00) for the CG. The test result [t(25) = –6.204, p = 0.000] proved that there was a significant difference in the scores of CG before and after the experiments. Therefore, it can be concluded that both of the two groups showed significant improvement as a result of the intervention.

The independent sample t-test analysis was used to compare the means of the EG and CG. Table 3 and Table 4 indicated descriptive statistics values for the variables in the pre-concept map and post-concept map scores respectively.

Table 4. Descriptive statistics for post-concept map result for VE & PE

	Lab type	N	Mean	Standard deviation	Standard error mean
PostCM	PE	26	44.00	12.133	2.376
	VE	25	42.40	5.339	1.068

Table 5. Independent sample t-test results for the pre- and post-concept map scores

Measurement	Group	N	M	SD	t	p*
Pre-concept map scores	Experimental	25	33.28	5.777	–1.204	0.235
	Control	26	31.00	5.010		
Post-concept map scores	Experimental	25	44.00	5.339	0.614	0.543
	Control	26	42.40	12.133		

* $p > 0.05$

Here, it can be interesting to look at each group's mean value. Table 3 and Table 4 demonstrate that the mean score of the EG increased from the pre-concept map score (M = 33.28, standard deviation [SD] = 5.777) to the post-concept map score (M = 42.40, SD = 5.339). Similarly, the descriptive statistics for the CG increased from the pre-concept map score (M = 31.00, SD = 4.501) to the post-concept map score (M = 44.00, SD = 12.113).

The descriptive statistics indicate that the mean score value of the CG is higher than that of the mean score of the EG in the post-concept map score. As can be seen in Table 4, students who practiced with physical experiments have slightly higher mean values than those who practiced with virtual experiments. As shown in Table 4, those who are instructed with physical experiments have a slightly higher mean value of 44.00 compared to 42.40 for those who are instructed with virtual experiments. This implies that, in chemical kinetics experiments, physical experiments are slightly better for understanding concepts than virtual experiments. Despite the descriptive statistics showing slight differences, it is important to use inferential statistics (independent sample t-test) to analyze significant differences between the two groups. The pre-concept map (Table 5) and post-concept map independent sample t-test analysis are shown for the two groups.

An independent-sample t-test was conducted to compare the pre-concept map and post-concept map scores for the control group (CG) and the experimental group (EG), as shown in Table 5. The test result [t(49) = –1.204, p = 0.235 > 0.05] proved that there was no significant difference in the scores of the control and experimental groups. Furthermore, the magnitude of the difference in means was calculated as mean difference = –1.742, with a 95% confidence interval of –4.649 to 1.166. Therefore, the two groups are similar before the intervention.

Post-concept map scores of the experimental and the control groups were also compared. The independent sample t-test compares the control (CG) and experimental (EG) groups whether there is a significant difference or not between the two groups in the post-concept maps (Table 5). Since the result showed that p > 0.05, the groups were not significantly different [t(df) = 0.614, p = 0.543 > 0.05] in the post-concept map scores. The magnitude of the differences in the mean (mean difference = 1.600), 95% confidence interval –3.689 to 6.889 was not significant. Hence the null hypothesis (H_0) was accepted.

For qualitative analysis, the propositions in the concept maps were analyzed for triplet representational levels such as macroscopic, microscopic, and symbolic levels applying a rubric in Table 2. This is demonstrated by students' phrasing in concept maps as the formation of a precipitate, bubbling of gases as macroscopic; a collision of reactive particles, forming and breaking of bonds as microscopic; chemical equations, rate equations as symbolic levels in their post-concept maps.

To provide a detailed example, a student's concept map from the control group was analyzed. In this map, the student used three types of representations but failed to establish cross-links between them. For instance, as shown in Table 6, the student illustrated the evolution of gas bubbles at the macroscopic level, noting a decrease in the mass of the system over time in the activity, demonstrating the reaction rate measurement. A symbolic representation was attempted with the equation $CaCO_3 + HCl \rightarrow CO_2$, although it was incomplete. Additionally, the student included a microscopic representation depicting the decomposition of calcium carbonate to produce gas, but this depiction was inaccurate. The concept map also indicated the formation of a white precipitate during the reaction between sodium thiosulfate and HCl, but the student struggled to explain the underlying microscopic processes.

In the experimental group, a student illustrated the impact of temperature on the reaction rate in the concept map. As shown in Table 6, the student explained that higher temperatures lead to an increase in molecular collisions, while lower temperatures result in fewer effective collisions. The student also emphasized the importance of sufficient energy to initiate the reaction. Additionally, he reported that higher temperatures result in the formation of more product molecules within a shorter time. The simulation, which included a stopwatch, further demonstrated that higher temperatures lead to faster production of products. When illustrating the effect of a catalyst, a student wrote "increase or decrease" on the connecting arrow, which does not accurately describe

Table 6. Examples of statements from qualitative analysis of students' concept maps

Qualitative analysis of students' concept maps
The concept map from the control group of higher scorer students demonstrates macroscopic representation, e.g., the formation of solid sulfur.
The concept map from the control group showed that a high intensity of gas bubbles was observed just a few seconds after the reaction started.
The formula for calculating the rate of reaction ($\Delta C/\Delta t$) was displayed, though the accompanying explanations were inaccurate in relating to the measurement of reaction rates (taken from the student in the experimental group).
Different reactions, such as combustion, fermentation, and fuel formation, were shown to have different rates (taken from the concept map of both student groups).
The requirement for a minimum amount of energy (activation energy) to start a reaction (taken from the student in the experimental group).
The concept of catalysts influencing reaction rates was presented, but the link between catalysts and reaction rates was not accurately represented (taken from the student in the experimental group).
The kinetic energy of molecules was shown to increase with temperature, but this idea was not well integrated with other concepts (taken from the student in the experimental group).
A white precipitate was observed to form quickly as the temperature of the reaction increased (taken from the concept map of student in the control group).
Overall, students in the control group appeared to remember more concepts and focused on macroscopic representations. However, students from the experimental group tended to use sub-microscopic representations. Despite this, the connections between concepts were shallow or inaccurately represented.

the effect of the catalyst on the reaction rate or reflect a meaningful relationship. These clues suggest that while microscopic representations were evident, macroscopic and symbolic representations were either missing or lacked connections or vice versa. Meaningful connections between triplet representations had not been successfully established in the middle and low-scoring students. Although relationships between concepts have been attempted, many were inaccurately expressed or lacked proper written explanations. Overall, although the number of concepts displayed increased in the post-concept map, the connections between these concepts were not effectively established.

DISCUSSION

The first objective of this study was to compare the conceptual understanding of students instructed with VEs and PEs applying concept maps as assessment tools to find out the effect of virtual experiments compared to physical experiments. The paired sample t-test result indicated that both types of laboratory activities showed significant improvement after the intervention on students' conceptual understanding. The independent sample t-test analysis for the post-concept maps indicated no significant difference between the groups, as demonstrated by the post-concept map scores of the students [$t(df) = 0.614$, $p = 0.543 > 0.05$], as indicated in **Table 5**. This means that virtual and physical laboratory activities are similar in enhancing conceptual learning. The result of this study is similar to a study that used concept maps as a pre- and post-assessment tool for computer-based instruction in the domain of acids and bases (Yaman and Ayas, 2015). The study stated that concept maps were effective in demonstrating changes in conceptual understanding in computer-based instruction.

Based on the findings, virtual experiments in chemical kinetics concepts are effective in the conceptual understanding of students. This is in agreement with Pyatt and Sims (2012) who claimed that variable inputs are easy to understand since VEs avoid distractions. According to Samon and Levy (2021), computer models and their appropriate pedagogies produce intended outcomes if used as recommended by scientific standards (National Research Council, 2012). Virtual laboratories are useful for schools with or without well-furnished physical laboratories. PE is often recommended for science learning, but VE can be most important for experiments that PE cannot reproduce in cases that involve hazardous chemicals (Serrano-Perez et al., 2023). Physical experiments help students understand concepts better, but virtual labs can also produce good outcomes not only for understanding concepts but also for science process skill mastery (Hunegnaw et al., 2024). Both virtual and physical laboratory activities help students improve their understanding of concepts. The findings of this research contradict another research report which claims that PE is less effective in helping students to write explanations or that VE performance better supports students' learning (Gnesdilow and Puntambekar, 2021). According to the aforementioned study, those students who used simulation were better at explaining and justifying the relationship between concepts. A significant decrease in understanding concepts and use of data was observed while students practiced in the physical lab.

The second research question in this study was to identify the most frequent triplet representation in the concept maps by small samples in both experimental and control groups from low, medium, and high scores of concept maps. The high-scoring student from the control group (PE) demonstrated a comprehensive understanding of the correlation between the three levels, exemplified by his ability to analyze precipitation formation (macroscopic nature) between sodium thiosulfate ($Na_2S_2O_3$) and that reacts with HCl from the concept map, which is attributed to the formation of sulfur (microscopic level). Furthermore, the student has represented the reaction through a chemical equation. This result is in line with Jaber and BouJaoude (2012) who indicated that carefully planned instruction could assist students in understanding the three representational levels (macroscopic, sub-microscopic, and symbolic).

However, the middle achievers and the high-profile student from the EG have demonstrated conceptual knowledge through symbolic and macroscopic levels, as indicated by the formation of gas bubbles (macroscopic level) during the reaction of $CaCO_3$ and HCl in the experimental group, which they were able to indicate through an equation, but with no clear evidence of the microscopic nature of the reaction. The low-scoring students' concept maps indicate the macroscopic representation but do not establish a link across the sub-microscopic representations. High school chemistry learning around the world requires students to master not only macroscopic or symbolic level concepts but also more abstract unobservable (microscopic) concepts (Hodges et al., 2018). Chemistry students throughout the world experience difficulty in connecting microscopic phenomena to their daily lives.

The qualitative analysis of this study indicated that students are better at showing their knowledge of macroscopic events than illustrating microscopic concepts. Despite most students improving in all representations, the macroscopic representation remained dominant. This is in agreement with Talanquer (2011) who states that many students fail to demonstrate the microscopic representations as the attribute is unobservable through sense organs. This finding is similar to other research on acid-base topics, where students represented their understanding at the macroscopic level due to their day-to-day experiences (Yaman and Ayas, 2015).

Limitations and Implications of the Study

Despite the rubric judging students' understanding could be imperfect, the information displayed by the concept maps is more important. Interpretation and scoring of concept maps pose a challenging and time-consuming process (Burry-Stock and Oxford, 1994). Teachers should have the skill to assess concept maps to facilitate students' learning. Allocating enough time and following up with all students for active involvement, including individual and group work and discussions, is crucial for successful concept map construction. Hsu and Hsieh (2005) reported that despite criticisms of heavy workload and time consumption, concept maps promote the development of problem-solving and critical-thinking skills through the organization of data and analysis of concept inter-linkages.

Concept maps can be made a powerful assessment tool since they relate the key nodes to the most important terms in the curriculum and develop a reliable scoring process. Repeated application of concept maps in assessment as well as instructions improves students' understanding (Vanides et al., 2005). Concept maps are very instructive, even without formal grading or scoring. Teachers could generate ideas for improving students' thinking/understanding from a shallow scan of the concept maps. Thus, one practical implication of this study is the gaps identified by concept maps could be used to modify lesson plans in classrooms for better learning.

Modern students rely extensively on computer technology, prompting educators to integrate it into classroom instruction to optimize learning outcomes (Hodges et al., 2018). The teaching of STEM subjects in Ethiopia also necessitates the use of updated virtual environments to achieve effective results. Technology enables successful chemistry experiments and elevates visualization in virtual laboratories. Additionally, educational games with a serious tone are available other than simulations to reinforce the understanding of microscopic chemistry concepts. This research recommends policy and curriculum designers include virtual experiments in Ethiopian science classrooms. Literature frequently promotes that combining virtual and physical experiments produces the most important outcome (Sypsas and Kalles, 2018; Wörner et al., 2022). The qualitative findings indicated that students struggle to interconnect the symbolic, microscopic, and macroscopic aspects of chemical kinetics. Addressing this gap (that is students' difficulty to represent the three representations-macroscopic, symbolic and sub-microscopic in a concept map) may require a combination of virtual and physical experiments. Therefore, future research may focus on the effective implementation of combining virtual and physical experiments in Ethiopian high schools, especially teaching microscopic levels to the macroscopic and symbolic levels of chemistry concepts.

CONCLUSION

This study examines the effectiveness of virtual experiments compared to physical experiments in the domain of chemical kinetics. To assess students' conceptual understanding, a more effective concept map assessment was utilized. Both qualitative and quantitative methods were employed to track students' progress and monitor changes in their conceptual knowledge. The quantitative analysis reveals that physical and virtual environments enhance conceptual learning as revealed from paired and independent sample t-tests. Virtual experiments, when guided with appropriate instruction and assessment, were effective in deepening high school students' conceptual understanding. Despite the quantitative analysis showing overall improvement in conceptual learning, the qualitative analysis of students' concept map representation was predominantly focused on macroscopic depictions. More specifically as revealed from the results from qualitative results, students struggled to connect the macroscopic, submicroscopic, and symbolic levels of representation. A better understanding of these levels could be achieved by combining virtual and physical experiments. Given the limited access to physical lab resources in many developing countries, the study's findings suggest that the concerned bodies (such as policymakers, curriculum designers, school administrators, and teachers) should consider incorporating more virtual experiments into broader science education programs.

ACKNOWLEDGEMENTS

This study was funded by the Federal Democratic Republic of Ethiopia Ministry of Education.

DECLARATION

The authors declare that there is no conflict of interest related to this work.

DATA AVAILABILITY STATEMENT

The data supporting the findings of this study will be available upon reasonable request from the corresponding author.

REFERENCES

Altmeyer, K., Kapp, S., Thees, M., Malone, S., Kuhn, J. and Brünken, R. (2020). The use of augmented reality to foster conceptual knowledge acquisition in STEM laboratory courses—Theoretical background and empirical results. *British Journal of Educational Technology*, 51(3), 611–628. https://doi.org/10.1111/bjet.12900

Ayalew, R. H. (2017). Quality of science teaching in secondary schools of North Gondar Zone in Ethiopia. *International Journal of Educational Sciences*, 16(1-3), 21–35. https://doi.org/10.1080/09751122.2017.1311592

Bekalo, S. and Welford, A. (1999). Secondary pre-service teacher education in Ethiopia: Its impact on teachers' competence and confidence to teach practical work in science. *International Journal of Science Education*, 21(12), 1293–1310. https://doi.org/10.1080/095006999290084

Bekalo, S. and Welford, G. (2000). Practical activity in Ethiopian secondary physical sciences: Implications for policy and practice of the match between the intended and implemented curriculum. *Research Papers in Education*, 15(2), 185–212. https://doi.org/10.1080/026715200402498

Belay, S., Atnafu, M., Michael, K. and Ermias, M. A. (2016). *Strategic policy for national science, technology and mathematics education*. Addis Ababa, Ethiopia: Ministry of Education. Available at: https://moe.gov.et/storage/Books/Strategic%20Policy%20for%20National%20Science%20Technology%20and%20Mathematics%20Education-Extracted.pdf.

Brinson. (2015). Learning outcome achievement in non-traditional (virtual and remote) versus traditional (hands-on) laboratories: A review of the empirical research. *Computers & Education*, 87, 218–237. https://doi.org/10.1016/j.compedu.2015.07.003

Burry-Stock, J. A. and Oxford, R. L. (1994). Expert science teaching educational evaluation model (ESTEEM): Measuring excellence in science teaching for professional development. *Journal of Personnel Evaluation in Education*, 8, 267–297. https://doi.org/10.1007/BF00973725

Cakmakci, G., Leach, J. and Donnelly, J. (2006). Students' ideas about reaction rate and its relationship with concentration or pressure. *International Journal of Science Education*, 28(15), 1795–1815. https://doi.org/10.1080/09500690600823490

Chan, P., Van Gerven, T., Dubois, J.-L. and Bernaerts, K. (2021). Virtual chemical laboratories: A systematic literature review of research, technologies and instructional design. *Computers and Education Open*, 2, 100053. https://doi.org/10.1016/j.caeo.2021.100053

Chiu, J. L., DeJaegher, C. J. and Chao, J. (2015). The effects of augmented virtual science laboratories on middle school students' understanding of gas properties. *Computers & Education*, 85, 59–73. https://doi.org/10.1016/j.compedu.2015.02.007

Cresswell, J. W. (2012). *Educational Research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed). Boston, MA: Pearson.

Domínguez, J., Miranda, R., González, E., Oliet, M. and Alonso, M. (2018). A virtual lab as a complement to traditional hands-on labs: Characterization of an alkaline electrolyzer for hydrogen production. *Education for Chemical Engineers*, 23, 7–17. https://doi.org/10.1016/j.ece.2018.03.002

Engida, T. (2002). Analysis of the science process skills in high school chemistry textbooks. *IER FLAMBEAU*, 10(1), 57–66.

Flegr, S., Kuhn, J. and Scheiter, K. (2023). When the whole is greater than the sum of its parts: Combining real and virtual experiments in science education. *Computers & Education*, 197, 104745. https://doi.org/10.1016/j.compedu.2023.104745

Gericke, N., Högström, P. and Wallin, J. (2023). A systematic review of research on laboratory work in secondary school. *Studies in Science Education*, 59(2), 245–285. https://doi.org/10.1080/03057267.2022.2090125

Gkitzia, V., Salta, K. and Tzougraki, C. (2011). Development and application of suitable criteria for the evaluation of chemical representations in school textbooks. *Chemistry Education Research and Practice*, 12(1), 5–14. https://doi.org/10.1039/C1RP90003J

Gnesdilow, D. and Puntambekar, S. (2021). Comparing middle school students' science explanations during physical and virtual laboratories. *Journal of Science Education and Technology*, 31, 191–202. https://doi.org/10.1007/s10956-021-09941-0

Hodges, G. W., Wang, L., Lee, J., Cohen, A. and Jang, Y. (2018). An exploratory study of blending the virtual world and the laboratory experience in secondary chemistry classrooms. *Computers & Education*, 122, 179–193. https://doi.org/10.1016/j.compedu.2018.03.003

Hsu, L. and Hsieh, S.-I. (2005). Concept maps as an assessment tool in a nursing course. *Journal of Professional Nursing*, 21(3), 141–149. https://doi.org/10.1016/j.profnurs.2005.04.006

Hunegnaw, T. and Melesse, S. (2023). An evaluative study of the experimental tasks of the Ethiopian grade 12 chemistry textbook considering developing "science process skills". *Cogent Education*, 10(1), 2208944. https://doi.org/10.1080/2331186X.2023.2208944

Hunegnaw, T., Hailegebreal, T. D., Getahun, D. A. and Atlabachew, M. (2024). Students' science process skills mastery through virtual experiments in chemical kinetics concepts. *Reflective Practice*. https://doi.org/10.1080/14623943.2024.2440174

Irby, B. J., Brown, G., Lara-Aiecio, R. and Jackson, S. (2013). *The Handbook of Educational Theories*. Charlotte, NC: Information Age Publishing.

Jaber, L. Z. and BouJaoude, S. (2012). A macro-micro-symbolic teaching to promote relational understanding of chemical reactions. *International Journal of Science Education*, 34(7), 973–998. https://doi.org/10.1080/09500693.2011.569959

Kapici, H. O., Akcay, H. and de Jong, T. (2019). Using hands-on and virtual laboratories alone or together—Which works better for acquiring knowledge and skills? *Journal of Science Education and Technology*, 28(3), 231–250. https://doi.org/10.1007/s10956-018-9762-0

Kaya, O. N. (2008). A student-centred approach: Assessing the changes in prospective science teachers' conceptual understanding by concept mapping in a general chemistry laboratory. *Research in Science Education*, 38, 91–110. https://doi.org/10.1007/s11165-007-9048-7

Krajcik, J. and Shin, N. (2023). Student conceptions, conceptual change, and learning progressions, in N. G. Lederman, D. L. Zeidler and J. S. Lederman (eds), *Handbook of Research on Science Education: Volume III*. Oxfordshire: Routledge. https://doi.org/10.4324/9780367855758-7

Krüger, J. T., Höffler, T. N., Wahl, M., Knickmeier, K. and Parchmann, I. (2022). Two comparative studies of computer simulations and experiments as learning tools in school and out-of-school education. *Instructional Science*, 50(2), 169–197. https://doi.org/10.1007/s11251-021-09566-1

Lopez, E., Kim, J., Nandagopal, K., Cardin, N., Shavelson, R. J. and Penn, J. H. (2011). Validating the use of concept-mapping as a diagnostic assessment tool in organic chemistry: Implications for teaching. *Chemistry Education Research and Practice*, 12(2), 133–141. https://doi.org/10.1039/C1RP90018H

Mayer, R. E. (2005). Cognitive theory of multimedia learning, in R. E. Mayer (ed), *The Cambridge Handbook of Multimedia Learning* (pp. 31–48). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.004

Millar, R. (2004). *The role of practical work in the teaching and learning of science*. Commissioned paper-Committee on High School Science Laboratories: Role and Vision. Washington DC: National Academy of Sciences, 308. Available at: https://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpage/dbasse_073330.pdf.

MoE, E. (2009). Curriculum framework for Ethiopian education (KG-Grade 12). *Addis Ababa*. Available at: https://moe.gov.et/storage/Books/Curriculum%20Framework%20for%20Ethiopian%20Education%20(KG%20%E2%80%93%20Grade%2012).pdf.

Nakhleh, M. B., Polles, J. and Malina, E. (2002). Learning chemistry in a laboratory environment, in J. K. Gilbert, O. De Jong, R. Justi, D. F. Treagust and J. H. Van Driel (eds), *Chemical Education: Towards research-based practice* (pp. 69–94). Heidelberg: Springer. https://doi.org/10.1007/0-306-47977-X_4

National Research Council. (2012). *A Framework for K-12 Science Education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press. https://doi.org/10.17226/13165

Nicoll, G. (2001). A three-tier system for assessing concept map links: A methodological study. *International Journal of Science Education*, 23(8), 863–875. https://doi.org/10.1080/09500690010025003

Penn, M. and Ramnarain, U. (2019). A comparative analysis of virtual and traditional laboratory chemistry learning. *Perspectives in Education*, 37(2), 80–97. https://doi.org/10.18820/2519593X/pie.v37i2.6

Puntambekar, S., Gnesdilow, D., Dornfeld Tissenbaum, C., Narayanan, N. H. and Rebello, N. S. (2021). Supporting middle school students' science talk: A comparison of physical and virtual labs. *Journal of Research in Science Teaching*, 58(3), 392–419. https://doi.org/10.1002/tea.21664

Pyatt, K. and Sims, R. (2012). Virtual and physical experimentation in inquiry-based science labs: Attitudes, performance and access. *Journal of Science Education and Technology*, 21, 133–147. https://doi.org/10.1007/s10956-011-9291-6

Rahmawati, Y., Hartanto, O., Falani, I. and Iriyadi, D. (2022). Students' conceptual understanding in chemistry learning using PhET Interactive Simulations. *Journal of Technology and Science Education*, 12(2), 303–326. https://doi.org/10.3926/jotse.1597

Ruiz-Primo, M. A. and Shavelson, R. J. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching*, 33(6), 569–600. https://doi.org/10.1002/(SICI)1098-2736(199608)33:6<569::AID-TEA1>3.0.CO;2-M

Samon, S. and Levy, S. T. (2021). The role of physical and computer-based experiences in learning science using a complex systems approach. *Science & Education*, 30, 717–753. https://doi.org/10.1007/s11191-020-00184-w

Schwendimann, B. A. (2015). Concept maps as versatile tools to integrate complex ideas: From kindergarten to higher and professional education. *Knowledge Management and E-Learning*, 7(1), 73–99. https://doi.org/10.34105/j.kmel.2015.07.006

Seery, M. K., Agustian, H. Y. and Zhang, X. (2019). A framework for learning in the chemistry laboratory. *Israel Journal of Chemistry*, 59(6–7), 546–553. https://doi.org/10.1002/ijch.201800093

Serrano-Perez, J. J., González-García, L., Flacco, N., Taberner-Cortés, A., García-Arnandis, I., Pérez-López, G., Pellin-Carcelen, A. and Romá-Mateo, C. (2023). Traditional vs. virtual laboratories in health sciences education. *Journal of Biological Education*, 57(1), 36–50. https://doi.org/10.1080/00219266.2021.1877776

Smeeth, L. and Ng, E. S.-W. (2002). Intraclass correlation coefficients for cluster randomized trials in primary care: Data from the MRC trial of the assessment and management of older people in the community. *Controlled Clinical Trials*, 23(4), 409–421. https://doi.org/10.1016/S0197-2456(02)00208-8

Sorden, S. D. (2012). The cognitive theory of multimedia learning, in B. J. Irby, G. Brown, R. Lara-Aiecio and S. Jackson (eds), *The Handbook of Educational Theories* (pp. 1–22). Charlotte, NC: Information Age Publishing.

Stoddart, T., Abrams, R., Gasper, E. and Canaday, D. (2000). Concept maps as assessment in science inquiry learning - A report of methodology. *International Journal of Science Education*, 22(12), 1221–1246. https://doi.org/10.1080/095006900750036235

Sypsas, A. and Kalles, D. (2018). Virtual laboratories in biology, biotechnology and chemistry education: A literature review, in *Proceedings of the 22nd Pan-Hellenic Conference on Informatics*, 29 November–1 December. Aigaleo: University of West Attica. https://doi.org/10.1145/3291533.3291560

Talanquer, V. (2011). Macro, submicro, and symbolic: The many faces of the chemistry "triplet". *International Journal of Science Edu*cation, 33(2), 179–195. https://doi.org/10.1080/09500690903386435

Tatli, Z. and Ayas, A. (2013). Effect of a virtual chemistry laboratory on students' achievement. *Journal of Educational Technology & Society*, 16(1), 159–170. https://www.jstor.org/stable/10.2307/jeductechsoci.16.1.159

Vanides, J., Yin, Y., Tomita, M. and Ruiz-Primo, M. A. (2005). Concept maps. *Science Scope*, 28(8), 27–31.

Watson, M. K., Pelkey, J., Noyes, C. R. and Rodgers, M. O. (2016). Assessing conceptual knowledge using three concept map scoring methods. *Journal of Engineering Education*, 105(1), 118–146. https://doi.org/10.1002/jee.20111

Winkelmann, K., Keeney-Kennicutt, W., Fowler, D., Lazo Macik, M., Perez Guarda, P. and Joan Ahlborn, C. (2020). Learning gains and attitudes of students performing chemistry experiments in an immersive virtual world. *Interactive Learning Environments*, 28(5), 620–634. https://doi.org/10.1080/10494820.2019.1696844

Wörner, S., Kuhn, J. and Scheiter, K. (2022). The best of two worlds: A systematic review on combining real and virtual experiments in science education. *Review of Educational Research*, 92(6), 911–952. https://doi.org/10.3102/00346543221079417

Yaman, F. and Ayas, A. (2015). Assessing changes in high school students' conceptual understanding through concept maps before and after the computer-based predict–observe–explain (CB-POE) tasks on acid–base chemistry at the secondary level. *Chemistry Education Research and Practice*, 16(4), 843–855. https://doi.org/10.1039/C5RP00088B

Zengele, A. G. and Alemayehu, B. (2016). The status of secondary school science laboratory activities for quality education in case of Wolaita Zone, Southern Ethiopia. *Journal of Education and Practice*, 7(31), 1–11.