Chemistry Education Research and Practice

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Chem. Educ. Res. Pract.*, 2026, **27**, 780

Virtual and hands-on laboratory environments in the chemistry classroom: the effect of prior chemistry knowledge

Michael A. Marino

With rapid technological advancements and the creation of innovative digital tools, virtual laboratory (VTL) experiences are being integrated into undergraduate chemistry curricula. Using a four-week counterbalanced design with 70 first-year college students, this study aimed to demonstrate how prior chemistry knowledge, assessed by a general chemistry exam, predicted student performance in traditional hands-on laboratory (HOL) and VTL experiments. The results of *t*-test analyses revealed similar overall performance between the HOL and VTL environments. However, the results of a MANOVA showed a differential effect: VTLs were more effective for students demonstrating a high level of prior chemistry knowledge, whereas no significant difference was observed for students with a low level of prior knowledge. Furthermore, prior knowledge explained 35% of the variance in high level knowledge students' VTL performance, indicating its strong predictive power for their laboratory success. These results suggest that although VTLs may not harm science achievement, they could inadvertently widen the performance gap between high-achieving and low-achieving students. Accordingly, this study underscores the importance of utilizing predictor variables, such as prior chemistry knowledge, to better understand and optimize the effectiveness of HOL and VTL experiments.

Received 8th August 2025,
Accepted 23rd January 2026

DOI: 10.1039/d5rp00293a

rsc.li/cerp

Virtual laboratory (VTL) experiences are rapidly being integrated into undergraduate chemistry curricula and experimentation (Brinson, 2015; Ali *et al.*, 2022). With advances in instructional technology, including simulations and virtual experimentation, chemistry instructors have reimagined laboratory education curricula to align laboratory activities with targeted learning outcomes and assessment (Chunwan and Greenberg, 2026).

Recent literature has revealed a debate involving the comparison (especially the effectiveness) of two laboratory learning environments: hands-on laboratory (HOL) and VTL experimentation (Brinson, 2017; Chan *et al.*, 2021; Elmoazen *et al.*, 2023). This ongoing discussion regarding student laboratory experiences is valuable because previous research has shown that HOL experiences play an essential role for students' learning science (Kontra *et al.*, 2015; Adkins, 2020; Hofstein and Hugerat, 2021). This shift is underscored by the fact that VTL experiences have garnered significant funding and increased support from national and international organizations to make science education accessible, sustainable, and globally connected (Royal Society of Chemistry, 2024).

School of Arts and Sciences, Molloy University, 1000 Hempstead Ave., Rockville Centre, NY 11571, USA. E-mail: mmarino1@molloy.edu

HOL *versus* VTL experimentation

Traditional laboratory curricula involve experiments that incorporate HOL investigations, utilizing chemicals, glassware, and other physical laboratory equipment to augment the scientific phenomena discussed in the lecture portion of the course. In contrast, the entire VTL experience is delivered on computers, tablets, or other mobile devices. For example, in a titration experiment designed by the Royal Society of Chemistry (2025), students viewed a video that provided the context of a real-world problem and then performed a virtual titration to calculate the acid concentration in an unknown stream sample. The simulation included realistic details, such as requiring students to tare the electronic balance after adding the weighing boat.

The virtual simulation experiments used in this study often involve software that places students in a simulated environment, incorporating both macro and micro dimensions through the use of images, interactive diagrams, and video, enabling students to visualize abstract concepts (*e.g.*, atoms moving at the molecular level; Chao *et al.*, 2016). VTLs offer advantages, such as cost-effectiveness, increased student engagement, and controlled environments where students can safely perform experiments without encountering the dangers associated with lab work (Brinson, 2017; Chan *et al.*, 2021, Ali *et al.*, 2022).

Review of the HOL *versus* VTL literature

Although there has been a major surge in VTL experiences in recent years, scholars have discussed the presence of computer-simulated VTLs in post-secondary science education as early as the 1970s (Campbell, 1985; Abdulwahed and Nagy, 2009). In the decades since, researchers have reported mixed results regarding the effectiveness of VTLs and performance measures (Ma and Nickerson, 2006; Chan *et al.*, 2021). Specifically in chemistry, several meta-analyses have been conducted to compare the effectiveness and variances between HOL and VTL experimentation (Brinson, 2015; Lynch and Ghergulescu, 2017). Although VTL experiments are practical for chemistry students, scholars collectively suggested that they should be considered as a complement to traditional HOL experimentation (Darrah *et al.*, 2014; Sypsas and Kalles, 2018). In addition, these findings underscore the need for further investigations into cost differences and preparation for advanced work in the field of chemistry (Faulconer *et al.*, 2018; Hofstein and Hugerat, 2021; Gungor *et al.*, 2022).

In several studies that found VTLs to be more effective, general trends emerged (Finkelstein *et al.*, 2005; Sypsas and Kalles, 2018; Chan *et al.*, 2021). First, these data indicated that students in the VTL were more thorough in their descriptions of the phenomenon, and second, that VTLs have led to more activities that facilitated higher-order thinking, as they allow students to engage in more metacognitive processes rather than focusing on laboratory techniques (Chan *et al.*, 2021). The effectiveness of VTL also varied depending on the teaching method. Compared to passive media, such as lectures or textbook readings, VTLs were more effective for improving conditional knowledge—that is, the metacognitive awareness of the conditions that influence learning and when they should be applied. Student motivation and attitudes also improved by creating a VTL-based lesson guided by a self-directed learning theoretical framework (Trúchly *et al.*, 2019), indicating a need for innovative pedagogical strategies to enhance the effectiveness of VTL experimentation (Reginald, 2023).

In studies that found no difference in the effectiveness of VTL and HOL experimentation, researchers suggested that VTLs adequately teach the content but may lead to a lack of proficiency in the manipulation of physical laboratory equipment (Zacharia and Constantinou, 2008; Tatli and Ayas, 2012, 2013; Chan *et al.*, 2021). For example, one study comparing HOL and VTL experimentation found no significant difference in the conceptual understanding of electrochemistry; however, a significant difference favoring the HOL group was observed in the physical placement of a salt bridge (Hawkins and Phelps, 2013). This finding aligns with the work of Merchant *et al.* (2014), who reported that the largest effect in VTL experiments was for declarative tasks ($d = 0.68$), rather than procedural tasks, such as operating the salt bridge ($d = 0.25$).

Theoretical framework

Cognitive load theory

This study was grounded in Cognitive Load Theory (CLT), a framework in education that explains how the limitations of working memory impact learning, particularly when students engage with complex tasks such as laboratory experimentation (Sweller, 1988). CLT categorizes three types of cognitive load: intrinsic load, which relates to the inherent difficulty of the content; extraneous load, which arises from the manner and structure of instruction; and germane load, the mental effort invested in constructing and refining mental models or schemas (Plass *et al.*, 2010). Effective instructional methods and pedagogy informed by CLT minimize extraneous load and optimize intrinsic load, freeing cognitive resources for germane processing to maximize learning outcomes (Sweller *et al.*, 2011; Lovell, 2020).

Application of CLT to HOL and VTL chemistry experiments

In chemistry laboratory learning, the application of CLT helps explain how students explore both HOL and VTL learning environments. For example, a highly complex laboratory procedure, coupled with factors such as proper safety protocols and the physical manipulation of glassware, can create significant extraneous cognitive load.

VTL experiments pose different sources of cognitive demand. For example, the investigation may be difficult to navigate, and VTLs can over-rely on passive observation, which may limit students with low prior knowledge. With this in mind, educators can design virtual environments to reduce extraneous load by streamlining procedures and focusing on core chemical principles, thereby enhancing germane load (Makransky *et al.*, 2019).

CLT's connection to prior knowledge

Research has identified prior knowledge as one of the most important factors influencing cognitive mode during learning (Mayer, 2001; Kalyuga, 2005). The only individual characteristic explicitly included in the CLT framework is students' prior knowledge (Kalyuga *et al.*, 1998). Since prior knowledge mediates cognitive load, a learner's existing knowledge influences the mental effort required to complete complex tasks in the laboratory. For example, when learners are novices (*i.e.*, students with low prior knowledge), the cognitive load associated with unguided discovery is too high to promote learning because novices lack well-developed schemas to guide their knowledge construction process (Tuovinen and Sweller, 1999; Kirschner *et al.*, 2006). In contrast, expert (high prior knowledge) students have an appropriate schema to build on their current knowledge with sufficient room in their working memory.

Students demonstrating a high level of prior chemistry knowledge may particularly benefit from the VTL setting, as the experiments and simulations focus on the theoretical aspects of chemical phenomena. By removing extraneous cognitive load, these students can focus on the topic of the experiment and effectively manage intrinsic cognitive load within their working memory capacity.

In contrast, students with low prior knowledge, who lack existing mental schemas, must deal with isolated elements of information and may rely on tactile experiences in the HOL environment. Since new information is cognitively demanding and affects working memory capacity, students with low prior knowledge levels may be disadvantaged when engaging with new material and curricula. Thus, CLT provides a fitting lens through which to view the context for investigating lab modality and potentially explains the difference in success in the high-level and low-level prior knowledge groups.

Prior chemistry knowledge as a predictor variable

Since researchers have identified prior knowledge as a key predictor of academic success, it was not surprising that numerous studies have supported the significant relationship between students' grade point average and academic achievement (Cassidy and Eachus, 2000; McKenzie and Schweitzer, 2001). These studies revealed that students' prior knowledge in any subject was the most important factor in predicting academic performance. Meta-analyses produced an effect size of 0.73 (*i.e.*, the magnitude of the difference between groups) for prior knowledge, which is considered relatively large based on Cohen's (1992) index of values (Sullivan and Feinn, 2012; Hattie, 2023). Furthermore, several studies have supported a significant relationship between prior knowledge and future science achievement, which is the basis of this study (Zeegers, 2004; McCall *et al.*, 2006).

The gap and research need

Since there is insufficient empirical evidence regarding moderator variables, such as prior chemistry knowledge, on the effectiveness of VTLs, the attribution of students' prior knowledge, perception, and attitudes would provide a more robust lens for evaluating the impact on future academic achievement in VTL settings and informing pedagogical methods and strategies to enhance student success. As comparative research between the two different laboratory environments continues to reveal the strengths and weaknesses of each, educators would benefit from measures to learn how best to utilize this technological advancement in conjunction with and/or instead of traditional labs.

Research questions and hypotheses

The research questions were derived from the growing conversation among chemistry educators seeking the most effective strategies for student learning and preparing students for future careers in science. The research questions are as follows:

1. To what extent does student performance differ in the VTL and HOL experiments?

2. To what extent does prior chemistry knowledge contribute to differences in student performance in VTL and HOL?

The first research question investigated students' ability to interact and succeed in the HOL and VTL settings, emphasizing appropriateness, accuracy, and quality of their responses to each question or prompt. The hypothesis was that students completing VTLs would outperform their HOL counterparts. The second research question investigated the extent to which prior chemistry knowledge, as measured by the participants' grades on a general chemistry assessment exam, predicts student academic success in HOL or VTL environments. The hypothesis was that post-secondary students with higher levels of prior chemistry knowledge would perform better in VTLs than in HOLs, while those with lower levels of prior chemistry knowledge would perform better in HOLs.

Methods

Participants

The 70 participants of this study were part of a convenience sample of students enrolled in an introductory level chemistry course titled "Organic and Biochemistry," designed specifically for nursing and other health-related majors at a four-year university on Long Island, New York. This was a single foundational course and did not include students concurrently enrolled in general chemistry or upper-level biochemistry courses. The data were collected in 2018 as part of the author's doctoral dissertation (Marino, 2018). The cohort primarily consisted of first-year college students (87.1%), was predominantly female (92.9%), and ranged in age from 18 to 22. All participants completed a high school general chemistry course within the past three years, which was the basis for the students' prior knowledge. IRB approval was obtained through a committee at the university prior to data collection, and included recruitment consent forms, emails to students, research design, data collection, and proposed analytics. All 70 students enrolled in the six-course sections provided the requisite consent. None of the students were later removed from the study; therefore, all 70 students were qualified for participation.

Research design

The study utilized a single-group counterbalanced A–B–A–B experimental design to investigate the influence of prior chemistry knowledge on success in VTL and HOL experiences. The students were assigned to Groups 1 and 2 based on their pre-registered course section and rotated through the four laboratory intervals in a counterbalanced A–B–A–B design (Table 1).

The experiment spanned approximately five instructional weeks. During the first week of the chemistry course, the students completed a 40-question "Prior Chemistry Assessment Exam" in lecture, and a 40-question "Pre-Test Assessment" in

Table 1 Treatment and control schematic for ABAB research design

		Interval			
Group	Pre-study	A	B	A	B
Group 1	Prior chemistry knowledge assessment	HOL	VTL	HOL	VTL
Group 2	Prior chemistry knowledge assessment	VTL	HOL	VTL	HOL

Note: Ns for groups 1 and 2 were 34 and 36, respectively. HOL = hands-on lab, control. VTL = virtual lab, treatment. Prior chemistry assessment was completed before the four instructional weeks in the lab. Post-tests were administered after every A and B interval.

the lab (both described in the Variable and Methods section and available in the SI document). In the next four weeks, the students performed experiments and completed a ''Post-Test Assessment'' – consisting of 10-matched items from the pre-test – at the end of each lab. Gain scores were calculated as the post-test score minus the pre-test score, using the 10-matched items corresponding to each lab, and gains were averaged across the four labs to provide a reasonable estimation of performance under each modality while reducing lab-specific nuances.

During the A (first and third) intervals, lesson content was delivered exclusively in the traditional HOL learning environment for the first group. These periods alternated with the instructional intervention of a VTL environment in the B (second and fourth) intervals. This model and intervals were flipped with the second group of participants of similar demographics. The VTL experiments, completed by the treatment intervals, originated from PhET Scientific. They were designed to be an identical laboratory experience as the HOL experiment, with the only changing variable being the mode in which students completed the experiment. For example, students connected atoms with bonds to create three-dimensional representations of organic molecules. Students in the HOL group connected balls representing atoms with wooden sticks representing bonds to create molecules. In contrast, those in the VTL group connected atoms while the program generated the three-dimensional shape according to VSEPR theory.

Regardless of participation, the four laboratory experiments and learning objectives were part of the department's traditional curriculum; therefore, the laboratory manual included four appendices with VTL versions of each of the four experiments. An overview of the topics, associated assessment item, and primary learning foci are provided in Table 2.

Irrespective of modality, the students were in the same laboratory, and the instructor presented a 30-minute pre-lab talk that provided (i) an overview, (ii) brief theory and calculations (if applicable), and (iii) proper safety and disposal measures. As students were completing the lab, the instructor walked around the room to answer safety and technological support questions. For their participation, the students were offered a chance to use their assessment scores as an exam grade in the course, thereby, providing a chance to increase their overall grade. However, if the students did not do well on the assessments, their performance had no bearing on their overall course grade. Additionally, if the students excelled in the post-tests, they were able to use their grade towards their homework grade.

Design rationale

The counterbalanced A–B–A–B design was chosen because it requires two baselines and two treatment phases. If scores revert to the baseline when the independent variable (*i.e.*, prior chemistry knowledge) is removed, and if intervention effects are similar during both treatment phases, the possibility that the effects are attributable to extraneous variables is significantly reduced (Gay *et al.*, 2012). Since a group was created for each possible order, the variance due to order effects became a separate source. The data for the labs were, therefore, not weighted. The counterbalanced design also allowed students to act as partial controls, providing an opportunity to experience both environments. Additionally, flipping modality aimed to mitigate systematic bias caused by practice effects and boredom effects and allows the analyses to focus on instructional modality effects across recurrent exposure.

At the end of the research period, IBM SPSS Statistics was used for data analysis. Prior knowledge was categorized to align with CLT as a categorical moderator rather than an exploratory *post hoc* grouping. This dichotomization ensured sufficient group sizing and aligned with the CLT educational framework, as three-way grouping would have reduced statistical power and inadvertently inflated variance. Although there were four different assessments, the combined gain scores represented an amassed measure of learning outcomes under the two instructional modalities across matched objectives. By combining gains across labs, the measurement reliability increases, and any error that may be due to the intricacies of a specific experiment is reduced.

Accordingly, *t*-test and MANOVA analyses evaluated whether the means of the dependent variables – achievement on the HOL or VTLs – were equal across levels of the categorical independent variable. MANOVA was appropriate for this analysis because it allowed the examination of multiple, correlated dependent variables and can be utilized to example how prior knowledge modulates responses across modalities. Given the homogenous demographic profile of the participants, gender and age were not considered covariates in the analysis. Assumptions of normality were assessed using Shapiro–Wilk tests, and

Table 2 Overview of laboratory topics and primary learning foci

Interval	Group 1 modality	Group 2 modality	Laboratory topic	Associated assessment items	Primary learning focus
Lab 1	HOL	VTL	Density	Items 1–10	Students calculated mass and volume relationships and differentiate between density values of various substances.
Lab 2	VTL	HOL	Acid–base chemistry	Items 11–20	Students differentiated between the pH values of acids, bases, and salts. They also calculated pH and observed indicators and buffers.
Lab 3	HOL	VTL	Molecular modelling	Items 21–30	Students connected atoms for molecular structure visualization and spatial reasoning.
Lab 4	VTL	HOL	Solubility	Items 31–40	Students observed the properties of solutions, investigated qualitative solubility rules, and examined conductivity.

Note: HOL = hands-on lab, control. VTL = virtual lab, treatment. Item numbers correlate with pre- and post-test assessment questions that are available on the SI document.

no substantial violations were observed. To explore specific variances, *post hoc* analyses using independent samples *t*-tests were used to investigate differences between the lab topic and modality. An analysis of covariance (ANCOVA) was not utilized, as no variances were attributed to covariates. The rejection level for all analyses was set at $p = 0.05$. Furthermore, Cohen's d effect sizes were obtained to provide a more profound understanding of how substantive the impact of the treatment was on the VTLs and how it varied among students of various levels of achievement.

Variables and measures

Prior chemistry knowledge

Participants took a general chemistry assessment to measure prior chemistry knowledge. The university has used the general chemistry assessment to measure students' comprehension upon entering the course and inform instructors on approaching pedagogical decisions and delivering curricula effectively. The students take the general chemistry assessment at the beginning of the course, regardless of their participation in the study, so these data were readily available.

The assessment was a 40-question multiple-choice exam modeled after the New York State (USA) Regents chemistry exam. It covered topics from general chemistry, such as atomic concepts, molar mass, and chemical bonding. Students took this exam during the first laboratory meeting to prevent confounding variables, such as the learning effect, where past experiences and study strategies significantly influence future learning and retention. The general chemistry assessment scores were used as a metric in the analysis to measure prior chemistry knowledge. Cronbach's alpha was calculated to evaluate internal consistency and reliability of the assessment exam. The analysis included all 40 items and yielded a Cronbach's alpha of 0.743. This value suggests that the assessment was sufficient and indicates a satisfactory degree to which the items were interrelated (Taber, 2018).

Pre-test assessment

The function of the pre-test assessment was to address the research questions and evaluate the students' knowledge before the treatments. The pre-test assessment was a 40-item multiple-choice quiz aligned with the topics and activities performed in the four lab experiments and yielded a Cronbach's alpha of 0.692, demonstrating acceptable reliability. Assessment questions focused on content-specific knowledge and not proper laboratory etiquette, equipment, and techniques, which were not formally covered in the VTL environment. The students had an hour to complete the assessment, and the researcher administered the exam. Each correct response received 1 point, and student scores were calculated out of 40 possible points. Pre-test assessment scores were compared with post-test assessment scores to calculate gain scores.

Post-test assessments

The function of the post-test assessments was to address the research questions and evaluate the students' knowledge after the treatment phases. The four post-test assessments were 10-item multiple-choice quizzes on the content of each laboratory experiment. Post-test question items were the same as the pre-test assessment questions, but the post-tests isolated the 10 questions that pertained to each appropriate lab. As in the pre-test, assessment questions focused on content-specific knowledge and not proper laboratory etiquette, equipment, and techniques. The post-test assessments were given to the treatment and control groups at the end of the laboratory period. The students had 20 minutes to complete the four assessments, and the researcher administered the exam. Each correct response received 1 point, and student scores were calculated out of 10. Post-test assessment scores were compared with pre-test assessment scores to create the students' gain score, which mitigated the possibility of practice effects from repeated exposure to the same test material.

Materials

PhET scientific laboratory simulations

PhET (Physics Education Technology), founded in 2002 by Carl Wieman, is an open educational resource project that offers free and effective science, technology, engineering, and mathematics (STEM) simulations. Researchers and faculty further developed this project at the University of Colorado–Boulder. The simulations are highly flexible and versatile learning tools that synergistically combine two-dimensional visuals, sound, and interactive features, such as the ability to increase the concentration of a solution to observe the change in pH and conductivity. PhET experiments have been translated into over 100 languages and display inclusive features, such as sound and sonification, alternative input, and interactive descriptions (Moore *et al.*, 2014).

This study used PhET simulations because they are well established, free, accessible to all learners, and widely regarded as an effective pedagogical tool. In addition, the simulations include research-based design that complements traditional curriculum and instruction at all levels of education. The topics of the four labs were density, acids and bases, solubility, and molecular modeling. The content of the labs featured interactive simulations of measuring mass and volume to calculate density, and immersing litmus paper and a conductivity meter into acidic and basic solutions, paralleling the HOL experimentation.

Results

Overall, the analyses show that HOL and VTLs were advantageous for learning. Prior chemistry knowledge predicted success for students entering the course with a high level of prior knowledge, but there was no difference for students with a low level of prior knowledge.

Prior chemistry knowledge

Based on their performance on the prior knowledge assessment, students were characterized into two groups: those demonstrating

Table 3 Descriptive statistics for prior chemistry knowledge, overall and by group

Variables	*n*	*M*	SD	Min	Max	Range	SE
Overall	70	22.95	3.50	12.0	37.0	25.0	0.42
High level of prior knowledge students	36	27.81	4.15	23.0	37.0	14.0	0.70
Low level of prior knowledge students	34	18.09	3.02	12.0	22.0	10.0	0.52

high levels of prior chemistry knowledge (scores $\geq$ 23, n = 36) and those demonstrating low levels of prior chemistry knowledge (scores $<$ 23, n = 34). The mean was used to determine the middle value for separating students into high- and low-knowledge groups, thereby ensuring a normal distribution of scores and sufficient sample sizes for statistical comparison. Aside from preserving statistical power and interpretability, dichotomization aligns with the CLT framework, as prior knowledge helps explain differences in academic achievement. Creating more than two groups would have reduced power, inflated variance, and made it difficult to observe interaction effects.

Table 4 Pre-, post-, and gain score comparison by prior knowledge group in HOL and VTL experiments

Lab topic	Modality	*n*	Pre-test *M* (SD)	Post-test *M* (SD)	Gain score *M* (SD)
High prior knowledge					
Lab 1 (density)	HOL	20	4.58 (2.83)	8.78 (1.20)	4.20 (0.74)
	VTL	16	5.45 (2.01)	8.56 (1.51)	3.11 (2.31)
Lab 2 (acids & bases)	HOL	16	6.00 (2.00)	8.22 (1.79)	2.22 (1.86)
	VTL	20	6.23 (2.80)	8.67 (1.12)	2.44 (2.19)
Lab 3 (solubility)	HOL	20	4.45 (1.88)	7.56 (1.01)	3.11 (1.69)
	VTL	16	3.89 (1.45)	7.67 (1.41)	3.78 (1.79)
Lab 4 (modeling)	HOL	16	4.44 (2.19)	7.22 (1.20)	2.78 (1.72)
	VTL	20	4.77 (1.45)	7.44 (.73)	2.67 (2.00)
Combined (All labs)	HOL	36	4.57 (1.95)	7.95 (1.67)	3.38 (1.95)
	VTL	36	4.93 (2.08)	8.09 (1.80)	3.16 (2.08)
Low prior knowledge					
Lab 1 (density)	HOL	16	4.80 (1.31)	8.00 (1.31)	3.20 (2.31)
	VTL	18	3.96 (1.99)	8.20 (.79)	4.24 (1.99)
Lab 2 (acids & bases)	HOL	18	4.88 (1.60)	7.79 (1.10)	2.91 (1.60)
	VTL	16	4.79 (1.60)	7.62 (1.19)	2.63 (1.60)
Lab 3 (solubility)	HOL	16	2.99 (1.31)	6.00 (2.39)	3.01 (1.31)
	VTL	18	3.56 (1.55)	6.80 (1.62)	3.24 (1.55)
Lab 4 (modeling)	HOL	18	3.00 (2.46)	7.40 (1.90)	4.40 (2.46)
	VTL	16	4.38 (1.60)	6.88 (1.46)	2.51 (1.60)
Combined (all labs)	HOL	34	4.22 (1.96)	7.30 (1.97)	3.08 (1.96)
	VTL	34	4.38 (2.09)	7.38 (2.04)	3.00 (2.09)

Note: HOL = Hands-on lab. VTL = Virtual lab. Gain = post–pre test scores for the ten items aligned with each laboratory experiment.

Descriptive statistics for the prior chemistry knowledge assessment, overall and by group, are presented in Table 3, which shows a clear separation between the high prior knowledge group (M = 27.81, SD = 4.15) and the low prior knowledge group (M = 18.09, SD = 3.02).

Gains score comparison on pre-test and post-test assessments

Pre-test, post-test, and gain scores were calculated for each group by prior knowledge and lab topic demonstrating high and low prior knowledge groups on all four HOL and VTL labs (Table 4). For each student, pre- and post-test comparisons were conducted using only the 10 items that were identical on both assessments. Frequency statistics were calculated for the eight HOL and VTL gain scores (post-test minus pre-test; see SI).

Correlations

Pearson correlation coefficients were calculated to examine the relationships: (i) within the same prior knowledge group across modalities, (ii) within the same modality across prior knowledge groups, and (iii) across modality and prior knowledge groups (Table 5). Correlation strengths were interpreted based on Cohen's (1992) conventions, where coefficients between 0.10 and 0.29 represent small associations and between 0.30 and 0.49 represent moderate associations. The interrelationships among lab gain scores were examined among lab gain scores to ensure suitability of multivariate analyses. For all scales, higher coefficients are indicative of stronger associations.

Small to moderate correlations were found among the dependent variables in all correlation matrices. Specifically, positive correlations were observed across most conditions; however, "mixed" directions were present, including a small negative correlation (r = −0.28) when comparing the VTL gains of high-prior knowledge students with the HOL gains of their low prior knowledge counterparts. These results confirm that the dependent variables are related but distinct, satisfying the requirements for multivariate analysis without concerns of multicollinearity. Most importantly, the combined lab gain scores showed stronger

Table 5 Selected Pearson correlations among combined lab gain scores by prior chemistry knowledge and modality

Category	Comp	Comparison	*r*
Within same prior knowledge group	HOL (High PK)	VTL (High PK)	0.33
	HOL (Low PK)	VTL (Low PK)	0.34*
Same modality across prior-knowledge groups	HOL (High PK)	HOL (Low PK)	0.49*
	VTL (High PK)	VTL (Low PK)	0.29
Cross modality, different prior knowledge	HOL (High PK)	VTL (Low PK)	0.24
	VTL (High PK)	HOL (Low PK)	−0.28

Note: HOL = hands-on laboratory experiments; VTL = virtual laboratory experiments; PK = prior knowledge. *$p < 0.01$.

correlations than individual labs, which emphasizes the two modalities and mitigates lab specific variance. Full correlation matrices of the four specific lab topics by modality and prior knowledge are reported in the Supplemental Information.

Research question 1: comparisons of HOL and VTLs. To test the hypothesis that VTL were more effective than HOL lab experiences, the scores from the treatment phases were compared with those from the control phases for each lab. A paired samples *t*-test was performed to compare the combined gain scores for the four combined HOLs and VTLs. Independent *t*-tests were then performed on each of the four HOLs and VTLs. For all five independent analyses comparing HOL and VTL experimentation, independent observations and normality assumptions were all met.

There was no significant difference between the combined hands-on and the combined virtual gain scores, $t(69) = 1.33$, $p = 0.19$, in the paired samples *t*-test, suggesting that total student performance did not differ between the HOL and VTL environments. There was no significant difference in gain scores for the four independent samples *t*-test analyses comparing each of the four HOL and VTLs, indicating that overall student performance did not differ between the two modalities (Table 6). Shapiro–Wilk's results were satisfactory for all four. Although none of the HOL *vs.* VTL comparisons reached statistical significance, the associated effect sizes ranged from negligible to small-to-moderate ($d = 0.01$–0.48), suggesting that the observed differences were generally small in magnitude and unlikely to be practically meaningful, even when considering the limited statistical power associated with the sample size.

Research question 2: the effect of prior chemistry knowledge on performance in HOLs and VTLs. A multivariate analysis of variance was carried out to assess the extent to which prior chemistry knowledge influenced student performance in the traditional HOL and VTL environments. The between-subjects factor comprised two groups: students who exhibited a high level of prior chemistry knowledge and students who exhibited a low level of prior chemistry knowledge. The dependent variables comprised scores on the HOL and VTL experiences. Evaluations of homogeneity of variance matrices and equality of variance were satisfactory, and moderate correlations were found among the dependent variables.

As shown in Table 7, the students who demonstrated a high level of prior chemistry knowledge scored significantly better than the students who demonstrated a low level of prior chemistry knowledge on the VTL experiences, $F(1,68) = 35.1$, $p = 0.001$, partial $\eta^2 = 0.35$; Shapiro–Wilk test, $W = 0.77$, $p = 0.14$. There was no significant difference between the two knowledge groups regarding HOL performance, $F(1,68) = 2.61$, $p = 0.11$, partial $\eta^2 = 0.04$. This indicates that 35% of the variance in VTL performance for the high prior knowledge students was explained by prior knowledge. The effect of prior knowledge on VTL performance produced a moderate effect size for students exhibiting high prior knowledge ($d = 0.42$) based on Cohen's (1992) convention. This value is a pairwise effect size from a *post hoc* analysis.

To test the hypothesis that VTL were as effective as HOL experiences for low prior knowledge students, gain scores were utilized to compare treatment and control groups for both lab environments. For both paired samples *t*-test analyses, independent observations and normality were all met. The high prior knowledge students scored significantly better on the VTL than the HOL post-tests, $t(35) = 2.49$, $p = 0.011$, $d = 0.42$; Shapiro–Wilk, $W = 0.81$, $p = 0.08$. However, there was no difference between HOL and VTL gains scores for the low prior

Table 6 Independent and paired samples *t*-tests comparing HOL and VTL gain scores

Comparisons	HOL M (SD)	VTL M (SD)	*t*(df)	*p*	Cohen's *d*
Lab experiments					
Lab 1: density (independent samples)	6.26 (2.81)	6.24 (3.05)	0.03 (68)	0.98	0.01
Lab 2: acids & bases (independent samples)	3.09 (1.99)	3.44 (1.97)	−0.86 (68)	0.40	−0.29
Lab 3: molecular modeling (independent samples)	2.62 (1.67)	2.26 (1.56)	0.90 (68)	0.37	0.30
Lab 4: solubility (independent samples)	3.38 (1.56)	4.00 (1.79)	−1.41 (68)	0.17	−0.48
Combined labs 1–4 (paired samples)	2.82 (1.98)	3.47 (2.09)	1.33 (69)	0.19	0.16

Note: Mean and standard deviation for gain scores, calculated as post-test minus pre-test performance for each lab. Independent samples *t*-tests were used to compare HOL and VTL gains scores for each individual lab. A paired samples *t*-tests was used to compare the combined HOL and VTL gain scores across all four laboratories (Shapiro–Wilk tests p's > 0.05 for all analyses).

Table 7 Summary of between-subjects effects and paired *post hoc* comparisons

Analysis type	Comparison	Test statistic	df	*p*	Effect size
Between subjects	High PK *vs.* low PK				
(Univariate ANOVAs)	HOL performance	$F = 2.61$	1, 68	0.110	$\eta^2 = 0.04$
	VTL performance	$F = 35.10$	1, 68	0.001	$\eta^2 = 0.35$
Within subjects	HOL *vs.* VTL modality				
(Paired *t*-tests)	High PK	$t = 2.49$	35	0.011	$d = 0.42$
	Low PK	$t = 0.70$	33	0.490	$d = 0.12$

Note: $N = 70$. High PK $n = 36$; low PK $n = 34$. PK = prior knowledge. η^2 = partial eta squared. All assumptions of normality (Shapiro–Wilk) and homogeneity were satisfied ($p > 0.05$).

knowledge students, $t(33) = 0.70$, $p = 0.49$, $d = 0.12$ (see Table 7). For this group, the VTL was found to be just as effective as the HOL environment.

Discussion

Summary of findings

The first research question of this study was designed to investigate possible differences in learning achievement in the HOL and VTL environments. The data revealed that students did not perform significantly better in either of these lab environments. Consequently, the results did not support the first hypothesis, which predicted that students completing VTLs would outperform their HOL counterparts. The literature supported this finding, as many studies found no difference in effectiveness between VTL and HOLs, suggesting that these environments can be used interchangeably in extenuating circumstances to save on costs of chemicals, glassware, and other materials (Zacharia and Constantinou, 2008; Tatli and Ayas, 2012, 2013; Darrah *et al.*, 2014; Chan *et al.*, 2021).

The second research question of this study investigated the extent to which prior chemistry knowledge influenced student learning in the HOL and VTL environments and how varying levels of prior knowledge impacted the effect of the intervention. The students with a high level of prior knowledge performed significantly better on the VTL experiments than the HOLs. Therefore, the second hypothesis was only partially supported. While the high prior knowledge students thrived in VTLs as predicted, the low prior knowledge students did not perform better in HOLs as expected. Instead, no significant differences were observed across lab modalities for this group.

Overall comparison of HOLs and VTLs

The positive effects of HOL and VTL experiments are complemented by the results of several studies and meta-analyses (*e.g.*, Chan *et al.*, 2021; Mercado and Picardal, 2023). These lines of evidence suggest that VTL may be used as a sensible supplement to the traditional HOL learning curriculum.

Considering the molecular modeling results, concepts like gas properties and collision theory may work better with VTLs so that students can see molecules and electrons move in a simulated environment. To illustrate, Mundy *et al.* (2024) demonstrated a design-based research approach to improve student learning by providing context and structure to scaffold abstract concepts with emission spectra. Like molecular shape and gas particles, light waves are another difficult concept for students, especially in a laboratory setting where hands-on approaches typically consolidate learning. Combined with HOL components, a virtually enhanced design can help preview physical lab experiences and boost collaboration, leading to higher student comfort and confidence in the overall environment.

Despite the many benefits of VTLs, chemistry educators have noted that VTLs alone may not effectively teach proper laboratory techniques and procedures (Hawkins and Phelps, 2013; Chan *et al.*, 2021). For example, the Royal Society of Chemistry (2023) emphasizes the importance of practical skills in developing students' understanding of chemical theory and phenomena. Although organizations acknowledge the usefulness and expediency of VTL experiments and simulations, the consensus is that there is no equivalent substitute for HOL experiences because they prepare students for highly technical careers and foster skills sought by potential employers.

The role of prior chemistry knowledge in laboratory performance

The study was designed to reframe assessment practices in the lab. Instead of measuring learning outcomes after the completion of the labs, this investigation focused on the effect of a predictor variable on the expected performance of students. The results of this study suggest that adding VTLs could benefit students entering the course with a low level of chemistry knowledge initially, but the effect was not as profound as that of their high-level counterparts. This finding challenges the initial prediction, based on CLT, that low prior knowledge students would struggle with the intricacies of VTL environments. One possible explanation is that the specific VTL approach provided sufficient scaffolding that benefited these students, particularly in relation to the learning objectives of the molecular modeling lab, an experiment that involves spatial reasoning. In this case, the VTL interface may have reduced the extraneous cognitive load associated with the physical manipulation in conventional HOLs, allowing students to concentrate on conceptual understanding rather than intricate laboratory technique. Conversely, for calculation-based objectives, the modality may have had a less pronounced effect on cognitive processing. In this case, educators should consider research that addresses the needs of low-performing students. Educators may also consider a more structured pre-laboratory design for students who struggle, such as an active learning platform for a large general chemistry course (Cresswell *et al.*, 2024). The active learning platform in the study was influenced by constructivist learning theories and included chances for self-assessment and revisions. Post-intervention data showed significant gains, specifically in the low prior knowledge student group.

While implementing effective, research-backed instructional strategies is vital for any educational endeavor, using predictor variables like prior chemistry knowledge can help forecast expected learning gains and achievement with the VTL experiences. Schools and universities would benefit from studying other key indicators, such as experience with technology, gaming experience, interest, and other motivational variables, to detect patterns and predict the efficacy of VTL experimentation. Chemistry departments that already use a pre-course and post-course assessment in their program can review the performance of the latest cohort of students to modify future laboratory curricula effectively.

Other considerations

The undergraduate students in this study showed no advantage from HOL or VTL modalities; however, previous research discovered that VTL experiences were more effective for adults than for children and adolescents, with the findings being statistically significant when comparing age groups (Muilwijk and Lazonder, 2023). Nevertheless, physical and virtual investigations yielded positive knowledge gains—specifically conceptual understanding—regardless of age, when touching materials provided chemistry students with relevant information about the lab's concepts. Instead, this phenomenon optimized the intrinsic load associated with the core learning of heat because the students feel the warmth while visually observing a temperature increase on their thermometer, thereby managing the inherent difficulty of the task and contributing to better learning within the CLT framework. On the other hand, VTL investigations were more effective when the touch experience was extraneous, particularly for abstract concepts like simulating the behavior of light rays in a spectrophotometer or when the focal variables, such as molecular movement, cannot be experienced by touch (Chan *et al.*, 2021; Mundy *et al.*, 2024).

Limitations

The present study involved 70 students enrolled in general chemistry, organic chemistry, and biological chemistry. The convenience sample consisted mainly of first-year and second-year undergraduate students majoring in chemistry, biology, environmental science, and nursing. Consequently, the experiment posed unique limitations.

The first limitation of this study was the sample size. Due to the nature of the convenience sample, the number of participants in the study was restricted. While the study's quasi-experimental A–B–A–B repeated-measures design allowed all participants to experience hands-on and virtual experimentation, smaller sample sizes can increase sensitivity to deviations from normality in statistical analyses. Still, all students received the same instruction, learned the same topics, and experienced new laboratory activities designed to fit the lab's learning and course objectives. Nevertheless, it's vital for future research to consider replicating the study with a larger number of participants to obtain a greater degree of separation between prior chemistry knowledge groups and include other STEM majors, such as engineering and computer science, to confirm the generalizability of the findings.

The second limitation was the timing of the experiment. The experiment was conducted over four instructional weeks, with several course sections. Therefore, students may have learned the corresponding lecture material throughout the time span. Although the learning objectives differed slightly from the lectures, the students still experienced the same topics each week in the lab. In other words, the students in the HOL and VTL completed the same topic in the same week, regardless of modality. For example, the students experimented with acids and bases in the second week of the laboratory investigation. In the lecture, one learning objective was for students to calculate the hydronium concentration in an aqueous solution. In contrast, in the lab, the students were asked to measure the pH with universal indicator paper and analyze the color to identify acidic, basic, and neutral compounds. The course instructors created properly scaffolded activities and learning goals with a pedagogical design to complement each other; however, mitigating the timing of instruction is difficult, considering the limited instructional days in a semester.

In conjunction with the first and second limitations, the third limitation was that no standardized assessments were used in this study. Each assessment needed to be specific and exclusive to the lab content for the experiment. Although they were similar to the New York State Regents Exam, the questions did not overlap completely. No existing standardized tests would have met this study's requirements, thereby creating the need to develop a new pre-test that students had to take in the course, regardless of their participation in the present study.

Fourth, the results provide promising evidence to support the implementation of VTLs for specific populations. However, additional studies are needed to generalize results beyond the specific contexts of this study and provide larger sample sizes. A reasonable direction for future research would be to conduct similar studies with participants from different geographic locations, including those with more ethnic or economic diversity, as it is possible that cultural and/or environmental factors in this specific population impacted the results of this study. Studies that examine the effects of gender and age group may produce interesting results because there have been mixed findings on affective domain outcomes, such as self-efficacy (Yu and Deng, 2022), and a few studies reported that gender does not positively or negatively affect performance on VTL assessments (Bazie *et al.*, 2024). Aside from gender, replication of this study with participants at the elementary and secondary levels would extend the results beyond the 18- to 20-year-old age range and provide more robust literature for the chemistry education community. The lab content, however, would need to be relevant to the age group.

Finally, only four lab topics were covered in this study, which limits the research findings due to the lack of content variety. Future research should include additional laboratory topics in general, such as organic or biochemistry, that students have never been exposed to, which would extend the findings of this research. In addition, high and low dichotomization provides a simplified, but theoretically aligned, operationalization within the CLT framework; however, those studies open the door for chemistry educators to conduct fruitful investigations that look at HOL and VTL experimentation. Regression-based moderation, with individual-level scores, would be a strong conceptual research design to provide further insight about how lab modality impacted learning gains on the different experiments and topics. Researchers could also use a regression model to target key predictor variables so that we can better decide curriculum design and delivery.

Implications and future direction

VTL learning has profound implications for science instruction and curriculum. The results of this study suggest that student performances in HOL and VTL environments did not differ; therefore, VTL experiments are an effective tool to consider as targeted supplement for instructional practices in chemistry. Although these data were collected in 2018, the findings are highly relevant to current curriculum challenges in science education, particularly given the increased incorporation of VTLs and simulations post-pandemic and ongoing concerns about equity in student learning.

The addition of VTL learning would affect various stakeholders, including students, teachers, schools, and curricula. Before procuring funds for large VTL programs, educators would benefit from reviewing the factors that may indicate success before investing in VTL instrumentation. Otherwise, administrators and curriculum designers should investigate other teaching methods, engaging programs, and affordable VTLs and simulations.

An additional implication of VTL integration simulations is accessibility and equity. Students can readily access VTLs on their tablet and mobile devices to review the content, which may help their understanding of the observed scientific phenomenon. Many simulations also contain accessible multimodal features such as interactive descriptions, sound, and sonification, which makes chemistry accessible to visually impaired students. These design features align with principles of CLT and the universal design for learning (UDL).

Due to the shift from a traditional teaching style to a more student-focused classroom—including an emphasis on inquiry and discovery—teachers should be trained on how to best use technology in the classroom. The shift to a student-focused classroom lends itself to emphasizing laboratory education since the laboratory provides a context to incorporate inquiry and discovery approaches naturally. Teachers would have the opportunity to create engaging curricula and supporting pedagogical activities. In addition, VTL experiments allow educators to create assessments and develop curricula to guide students toward a well-connected series of learning goals.

In this experiment, each lab concept was introduced to the students in the previous lecture class to ensure that the students were exposed to the materials. In science education, teaching the content in the lecture and then reinforcing it in the lab is traditional practice. For example, if students had just completed a lesson on solubility, the following lab class would be dedicated to reinforcing the new concepts using HOL or VTL techniques. Not only does this follow-up provide relevant connections to student learning objectives, but it also emphasizes student accountability for their independent learning. In response to new research-based best practices in STEM, newer instructional methods and learning theories (*e.g.*, flipped classroom and universal design for learning), have appropriately challenged the traditional lecture as an ineffective passive-learning method (*e.g.*, Kapur *et al.*, 2022). With the development of newer technologies and virtual experiences for students, educators can harness the power of VTL simulations to help students match the macroscopic scale with the atomic level. Technology, VTL experiments, and simulations are more effective for consolidation—rather than encoding—when there are multiple learning opportunities (Hattie, 2023). Therefore, repeated exposure to content through VTL simulations can aid students' long-term memory storage and retrieval on summative assessments.

Conclusions

College undergraduates are at a unique learning stage as they prepare for their future careers in science, engineering, and medicine that will rely more on technologies to perform routine tasks (*e.g.*, building structures and aiding in complex surgical procedures). The evolving landscape of science may make VTLs increasingly relevant tools, as these environments prepare students for future careers that rely on technology. In this context, educators have the opportunity to develop a rigorous instructional design that uses well-established learning theories and instructional support to optimize the effectiveness of HOL and VTL experiments by leveraging the positive features of each learning environment. In VTL simulations, students can manipulate constants like temperature or pressure and examine the outcomes in an engaging environment. Students can observe bubbles, steam, and sounds on the HOL macroscopic scale and witness increased collisions on the VTL atomic level. The amalgamation of the dual modalities can help students analyze causal relationships and inspire educators to create innovative and stimulating pedagogy regardless of prior knowledge or ability.

This study's results suggest that VTLs do not harm science achievement and may conserve resources. However, they may differentially impact students depending on prior knowledge. Although previous research has illuminated the strengths and weaknesses of HOL and VTL experimentation, this study has further addressed critical gaps regarding the advantageous nature of predictor variables in foreseeing the effectiveness of HOL and VTL experiments. While this initial investigation yields an interesting finding, further research is necessary to fully explore the broader significance of the interaction between prior knowledge and modality. Prospective studies may replicate findings and investigate the effects of other predictor variables, aside from prior chemistry knowledge, on the various types of VTL experiences. These insights may inspire chemistry educators to develop innovative curricula alongside pioneering technological advancements to better support all students.

Conflicts of interest

There are no conflicts to declare.

Data availability

The datasets generated and analyzed during this study, collected in 2018 as part of the author's doctoral dissertation, have been included in the Results section, the supplementary information (SI) document, and below. While the aggregate descriptive statistics are provided for all analyses, the original individual assessment files are no longer available due to the time elapsed since data collection and institutional data retention policies. All relevant summary statistics and derived values essential for replicating the report analyses are included. The supplementary information includes Survey Instrument 1 (Prior Chemistry Knowledge Assessment), Survey Instrument 2 (Pre-test and Post-test Assessment by Topic), Pre-test and Post-test Assessments and Gain

Scores, and Pearson Correlations. See DOI: https://doi.org/10.1039/d5rp00293a.

Acknowledgements

The author would like to acknowledge Dr. Bruce Torff, Professor of Education and graduate research mentor from Hofstra University, USA, for his research insights, countless hours of editing, and, most importantly, inspiration to produce high-quality research to better the field of chemical education. The author would also like to acknowledge Dr. Ling Huang, Professor of Chemistry and undergraduate research mentor from Hofstra University, USA, for his support, encouragement, and phenomenal lessons on chemical laboratory techniques and instrumentation. Last, the author would like to acknowledge Dr. Amy Catalano, Professor of Education, from Hofstra University, USA, for her wisdom on succinct writing, detailed data analysis, and the art of literature review.

References

Abdulwahed, M. and Nagy, Z. K., (2009), Applying Kolb's experiential learning cycle for laboratory education, *J. Eng. Educ.*, **98**, 283–294, DOI: 10.1002/j.2168-9830.2009.tb01025.x.

Adkins, D. G., (2020), Effects of hands-on experiences on student achievement, interest, and attitude in chemistry [Doctoral dissertation, Stephen F. Austin State University], https://scholarworks.sfasu.edu/etds/343.

Ali, N., Ullah, S. and Khan, D., (2022), Interactive laboratories for science education: A subjective study and systematic literature review, *Multimodal Technol. Interact.*, **6**(10), 85, DOI: 10.3390/mti6100085.

Bazie, H., Lemma, B., Workneh, A. and Estifanos, A., (2024), The effect of virtual laboratories on the academic ahievement of undergraduate chemistry students: Quasi-experimental study, *JMIR Form Res.*, **8**, e64476, DOI: 10.2196/64476.

Brinson, J. R., (2015), Learning outcome achievement in non-traditional (virtual and remote) versus traditional (hands-on) laboratories: A review of the empirical research, *Comput. Educ.*, **87**, 218–237, DOI: 10.1016/j.compedu.2015.07.003.

Brinson, J. R., (2017), A further characterization of empirical research related to learning outcome achievement in remote and virtual science labs, *J. Sci. Educ. Technol.*, **26**(5), 546–560, DOI: 10.1007/s10956-017-9699-8.

Campbell, D. R., (1985), Interactive graphics software for an undergraduate course in digital signal processing, *Comput. Educ.*, **9**(2), 79–86, DOI: 10.1016/0360-1315(85)90029-6.

Cassidy, S. and Eachus, P., (2000), Learning style, academic belief systems, self-report student proficiency and academic achievement in higher education, *Educ. Psychol.*, **20**(3), 307–320, DOI: 10.1080/713663740.

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, *Comput. Educ.*, **2**, 100053, DOI: 10.1016/j.caeo.2021.100053.

Chao, J., Chiu, J., DeJaegher, C. and Pan, E., (2016), Sensor-augmented virtual labs: Using physical interactions with science simulations to promote understanding of gas behavior, *J. Sci. Educ. Technol.*, **25**(1), 16–33, DOI: 10.1007/s10956-015-9574-4.

Chunwan and Greenberg, S., (2026), Lowering the activation barrier to create a constructively aligned undergraduate chemistry laboratory experience – a review of innovations in assessments and course design, *Chem. Educ. Res. Pract.*, **27**, 13–30, DOI: 10.1039/d5rp00050e.

Cohen, J., (1992), A power primer, *Psychol. Bull.*, **112**(1), 155–159, DOI: 10.1037//0033-2909.112.1.155.

Cresswell, S. L., Loughlin, W. A. and Kim, T. H., (2024), Implementing an interactive online platform in a large undergraduate general chemistry course and its impact on student learning and perceptions, *Chem. Educ. Res. Pract.*, **25**, 703–720, DOI: 10.1039/D3RP00221G.

Darrah, M., Humbert, R., Finstein, J., Simon, M. and Hopkins, J., (2014), Are virtual labs as effective as hands-on labs for undergraduate physics? A comparative study at two major universities, *J. Sci. Educ. Technol.*, **23**(6), 803–814, DOI: 10.1007/s10956-014-9513-9.

Elmoazen, R., Saqr, M. and Khalil, M., (2023), Learning analytics in virtual laboratories: A systematic literature review of empirical research, *Smart Learn. Environ.*, **10**, 23, DOI: 10.1186/s40561-023-00244-y.

Faulconer, E. K., Griffith, J. C., Wood, B. L., Acharyya, S. and Roberts, D. L., (2018), A comparison of online and traditional chemistry lecture and lab, *Chem. Educ. Res. Pract.*, **19**, 392–397, DOI: 10.1039/C7RP00173H.

Finkelstein, N. D., Adams, W. K., Keller, C. J., Kohl, P. B., Perkins, K. K., Podolefsky, N. S., Reid, S. and LeMaster, R., (2005), When learning about the real world is better done virtually: A study of substituting computer simulations for laboratory equipment, *Phys. Rev. ST Phys. Educ. Res.*, **1**, 010103, DOI: 10.1103/PhysRevSTPER.1.010103.

Gay, L. R., Mills, G. E. and Airasian, P., (2012), *Educational research: Competencies for analysis and applications*, 10th edn, Pearson Education, Inc.

Gungor, A., Kool, D., Lee, M., Avraamidou, L., Eisink, N., Albada, B., van der Kolk, K., Tromp, M. and Bitter, J. H., (2022), The use of virtual reality in a chemistry lab and its impact on students' self-efficacy, interest, self-concept and laboratory anxiety, *Eurasia J. Maths., Sci. Technol. Educ.*, **18**(3), em2090, DOI: 10.29333/ejmste/11814.

Hattie, J., (2023), *Visible learning, the sequel: A synthesis of over 2100 meta-analyses relating to achievement*, 2nd edn, Routledge.

Hawkins, I. and Phelps, A. J., (2013), Virtual laboratory vs. traditional laboratory: Which is more effective for teaching electrochemistry? *Chem. Educ. Res. Pract.*, **14**, 516–523, DOI: 10.1039/C3RP00070B.

Hofstein, A. and Hugerat, M., (2021), *Teaching and learning in the school chemistry laboratory*, The Royal Society of Chemistry.

Kalyuga, S., (2005), Prior knowledge principle in multimedia learning, In R. Mayer (ed.), *Cabridge handbook of multimedia* *learning* (pp. 235–337), New York: Cambridge University Press.

Kalyuga, S., Chandler, P. and Sweller, J., (1998), Expertise reversal effect, *Educ. Psychol.*, **38**, 23–31.

Kapur, M., Hattie, J., Grossman, I. and Sinha, T., (2022), Fail, flip, fix, and feed – rethinking flipped learning: A review of meta-analyses and a subsequent meta-analysis, *Front. Educ.*, **7**, 956416, DOI: 10.3389/feduc.2022.956416.

Kirschner, P., Sweller, J. and Clark, R., (2006), Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential and inquiry-based teaching, *Educ. Psychol.*, **41**, 75–86.

Kontra, C., Lyons, D. J., Fischer, S. M. and Beilock, S. L., (2015), Physical experience enhances science learning, *Psychol. Sci.*, **26**(6), 737–749, DOI: 10.1177/0956797615569355Lo.

Lovell, O., (2020), *Sweller's cognitive load theory in action*, 1st edn, John Catt Educational.

Lynch, T. and Ghergulescu, I., (2017), Review of virtual labs as the emerging technologies for teaching STEM subjects, Proceedings of the 11th International Technology, Education and Development Conference, 6082–6091, DOI: 10.21125/inted.2017.1422.

Ma, J. and Nickerson, J. V., (2006), Hands-on, simulated, and remote laboratories: A comparative literature review, *ACM Comput. Surv.*, **38**(3), 1–24, DOI: 10.1145/1132960.1132961.

Makransky G., Borre-Gude S. and Mayer, R. E., (2019), Motivational and cognitive benefits of training in immersive virtual reality based on multiple assessments, *J. Comput.-Assisted Learn.*, **35**(6), 691–707, DOI: 10.1111/jcal.12375.

Marino, M. A., (2018), Virtual and hands-on laboratory environments in the science classroom: The effect of prior science achievement (ED591485), Hofstra University.

Mayer, R. E., (2001), *Multimedia learning*, New York: Cambridge University Press.

McCall, K. L., Allen, D. D. and Fike, D. S., (2006), Predictors of academic success in a doctor of pharmacy program, *Am. J. Pharm. Educ.*, **70**(5), 106, DOI: 10.5688/aj7005106.

McKenzie, K. and Schweitzer, R., (2001), Who succeeds at university? Factors predicting academic performance in first year Australian university students, *Higher Edu. Res. Dev.*, **20**(1), 21–33, DOI: 10.1080/07924360120043621.

Mercado, J. C. and Picardal, J. P., (2023), Virtual laboratory simulations in biotechnology: A systematic review, *Sci. Educ. Int.*, **34**(1), 52–57, DOI: 10.33828/sei.v34.i1.6.

Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W. and Davis, T. J., (2014), Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis, *Comput. Educ.*, **70**, 29–40, DOI: 10.1016/j.compedu.2013.07.033.

Moore, E. B., Chamberlain, J. M., Parson, R. and Perkins, K. K., (2014), PhET interactive simulations: Transformative tools for teaching chemistry, *J. Chem. Educ.*, **91**(8), 1191–1197, DOI: 10.1021/ed4005084.

Muilwijk S. E. and Lazonder A. W., (2023), Learning from physical and virtual investigation: A meta-analysis of conceptual knowledge acquisition, *Front. Educ.*, **8**, 1163024, DOI: 10.3389/feduc.2023.1163024.

Mundy, C. E., Potgieter, M. and Seery, M. K., (2024), A design-based approach to improving pedagogy in the teaching laboratory, *Chem. Educ. Res. Pract.*, **25**, 266–275, DOI: 10.1039/D3RP00134B.

Plass, J. L., Moreno, R. and Brunken, R. (ed.) (2010), *Cognitive Load Theory*, Cambridge University Press.

Reginald, G., (2023), Teaching and learning using virtual labs: Investigating the effects on students' self-regulation, *Cogent Educ.*, **10**(1), 2172308, DOI: 10.1080/2331186X.2023.2172308.

Royal Society of Chemistry, (2023), *Position statement: Practical chemistry education: A policy position on the Royal Society's vision for practical chemistry in 5-19 education*, Retrieved from May 2, 2025, from https://www.rsc.org/globalassets/22-new-perspectives/education/01-curriculum-for-all/royal-society-of-chemistry—our-policy-position-on-practical-chemistry-education.pdf.

Royal Society of Chemistry, (2024), *Digital virtual reality centres for global equality in scientific training: The 3DI virtual reality institute – 2024 horizon prize for education winner*, Retrieved from June 2, 2025, from https://www.rsc.org/prizes-funding/prizes/2024-winners/the-3di-virtual-reality-institute/.

Royal Society of Chemistry, (2025), *Titration screen experiment*, Retrieved June 12, 2025, from https://virtual.edu.rsc.org/titration/experiment/2.

Sullivan, G. M. and Feinn, R., (2012), Using effect size—or why the *p* value is not enough, *J. Grad. Med. Educ.*, **4**(3), 279–282, DOI: 10.4300/JGME-D-12-00156.1.

Sweller, J., (1988), Cognitive load during problem solving: Effects on learning, *Cognit. Sci.*, **12**(2), 257–285, DOI: 10.1207/s15516709cog1202_4.

Sweller, J., Ayres, P. and Kalyuga, S., (2011) *Cognitive Load Theory*, New York: Springer.

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, Association for Computing Machinery, New York, pp. 70–75, DOI: 10.1145/3291533.3291560.

Taber, K. S., (2018), The use of Cronbach's alpha when developing and reporting research instruments in science education, *Res. Sci. Educ.*, **48**, 1273–1296, DOI: 10.1007/s11165-016-9602-2.

Tatli, Z. and Ayas, A., (2012), Virtual chemistry laboratory: Effect of constructivist learning environment, *Turk. Online J. Distance Educ.*, **13**, 183–199, Retrieved May 16, 2025, from https://eric.ed.gov/?id=EJ976940.

Tatli, Z. and Ayas, A., (2013), Effect of a virtual chemistry laboratory on students' achievement, *J. Educ. Technol. Soc.*, **16**, 159–170, Retrieved May 16, 2025, from https://eric.ed.gov/?id=EJ1016363.

Trúchly, P., Medvecký, M., Podhradský, P. and Mawas, N., (2019), STEM education supported by virtual laboratory incorporated in self-directed learning process, *J. Electr. Eng.*, **70**(4), 332–344, DOI: 10.2478/jee-2019-0065.

Tuovinen, J. and Sweller, J., (1999), A comparison of cognitive load associated with discovery learning and worked examples, *J. Educ. Psychol.*, **91**, 334–341.

Yu, Z. and Deng, X., (2022), A meta-analysis of gender differences in e-learners' self-efficacy, satisfaction, motivation, attitude, and performance across the world, *Front. Psychol.*, **13**, 2018–2025, DOI: 10.3389/fpsyg.2022.897327.

Zacharia, Z. C. and Constantinou, C. P., (2008), Comparing the influence of physical and virtual manipulatives in the context of the physics by inquiry curriculum: The case of undergraduate students' conceptual understanding of heat and temperature, *Am. J. Phys.*, **76**(4), 425–430, DOI: 10.1119/1.2885059.

Zeegers, P., (2004), Student learning in higher education: A path analysis of academic achievement in science, *Higher Educ. Res. Dev.*, **23**(1), 35–56, DOI: 10.1080/0729436032000168487.