



Virtual Laboratories in Undergraduate Science and Engineering Courses: a Systematic Review, 2009–2019

Shalaunda M. Reeves¹ · Kent J. Crippen²

Accepted: 16 September 2020 / Published online: 9 October 2020
© Springer Nature B.V. 2020

Abstract

Current technology has the capacity for affording a virtual experience that challenges the notion of a teaching laboratory for undergraduate science and engineering students. Though the potential of virtual laboratories (V-Labs) has been extolled and investigated in a number of areas, this research has not been synthesized for the context of undergraduate science and engineering education. This study involved a systematic review and synthesis of 25 peer-reviewed empirical research papers published between 2009 and 2019 that focused on V-Labs. The results reveal a dearth of varied theoretical and methodological approaches where studies have principally been evaluative and narrowly focused on individual changes in content knowledge. The majority of studies fell within the general domain of science and involved a single 2D experience using software that was acquired from a range of outside vendors. The perspective largely assumed V-Labs to be a teaching approach, providing instruction without any human-to-human interaction. Positive outcomes were attributed, based more on novelty than design, to improved student motivation. Studies exploring individual experiences, the role of personal characteristics or environments that afforded social learning, including interactions with faculty or teaching assistants were noticeably missing.

Keywords Virtual laboratories · Undergraduate · Science education · Engineering education · Laboratory education

Introduction

A laboratory course experience is recognized as a key component and degree requirement for most undergraduate science and engineering majors and has generally been considered the backbone of science education (Hofstein and Lunetta 2004; Reid and Shah 2007). Such experiences, which typically occur in specially designated physical spaces known as teaching laboratories, are intended to provide students with the opportunity to put theory into practice in the form of appropriate experiments at a given level of advancement in a discipline or with specific topics within a course or program of study (Ural 2016). The activities that occur in such courses are typically designed as expository experiences that both

require students to follow specified procedures such as performing experiments or making observations while at the same time demonstrating skills and understanding of associated concepts (Kempton et al. 2017). This implies that current laboratory education lacks activity akin to that of a practicing scientist, which would include “determin[ing] the problems, develop[ing] solutions and alternative solutions for these problems, search[ing] for information, evaluat[ing] the information and communicat[ing] with [peers]” (Ural 2016 p. 217). In essence, a teaching laboratory is more likely to consist of rote procedural tasks that do not support the development of critical thinking, ones that could be used to translate concepts, theories, or arguments beyond a superficial level of knowing (Basey et al. 2000; Gobaw 2016; Ryker and McConnell 2017; Wan et al. 2020).

Current technology has the capacity to challenge the traditional notion of a teaching laboratory and subsequently our perspective on the required laboratory experience for undergraduate students (Bernhard 2018). Virtual laboratories (V-Labs) are technology-mediated experiences in either two- or three-dimensions that situate the student as being in an emulation of the physical laboratory with the capacity to manipulate virtual equipment and materials via the keyboard and/or handheld controllers. Such experiences are typically delivered

✉ Kent J. Crippen
kcrippen@coe.ufl.edu

Shalaunda M. Reeves
smreeves@umich.edu

¹ Center for Academic Innovation, University of Michigan, Ann Arbor, MI, USA

² School of Teaching and Learning, University of Florida, PO Box 117048, Gainesville, FL 32611, USA

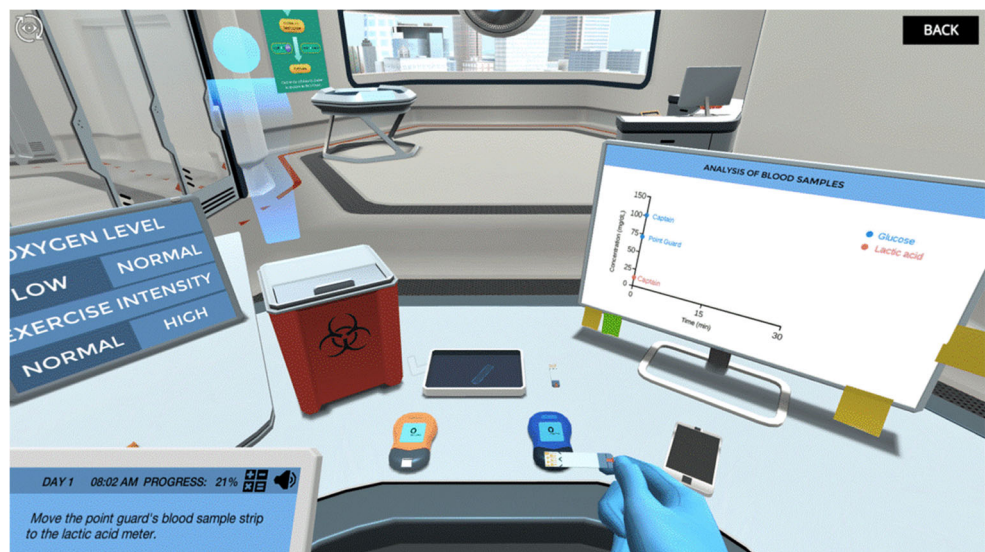
in two-dimensions (2D) with either a desktop or laptop computer, which is considered a low-immersion technology, or in three-dimensions (3D) with a head-mounted display, which is classified as a high-immersion technology. The use of the term immersion thus describes these experiences as a measure of the technology used and implies a collection of affordances for the student (Cummings and Bailenson 2016). The actual experience of the student, often captured as the variable presence or feeling of being in the emulated environment, involves the interplay of their psychology in response to these affordances, which includes their sensory perception, level of control, and ability to modify the environment (Riva et al. 2003).

V-Labs typically entail making observations and completing experiments that may involve the testing of theories and/or hypotheses (de Jong et al. 2014; Potkonjak et al. 2016). As such, this technology has the potential for delivering a first-person experience that very closely approximates not just that of a teaching laboratory (Vrellis et al. 2016) but that of a research laboratory, a situation where the student stands virtually in the shoes of a researcher and performs complex research skills as part of a larger multi-part task (Makransky et al. 2016, 2017) (Fig. 1). Examples of such skills would include preparing solutions, pipetting or creating calibration curves, as well as setting up and operating specialized analytical equipment such as a spectrophotometer or PCR machine. To date, evaluation studies in certain STEM contexts have shown that student achievement or performance for such V-Labs is consistent with that found in traditional face-to-face laboratory experiences (Darrah et al. 2014; Ekmekci and Gulacar 2015; Goudsouzian et al. 2018; Hawkins and Phelps 2013; Koh et al. 2010; Makransky et al. 2016; Ogbuanya and Onele 2018; Olympiou and Zacharia 2012; Vrellis et al. 2016).

A V-Labs' contextually rich environment and high-fidelity approximation of research practice is hypothesized to provide students with transferable knowledge (Baker et al. 2016) in the service of adaptive expertise (Alexander 2003). Transferable knowledge is that which has been acquired in one situation, but is available and can be used in the performance of a new or novel task in another situation or context (Kester et al. 2001). If students perceive and experience a V-Lab as if physically being in that laboratory, then their proficiency with any skills they acquire should translate to future work in a physical laboratory (Jensen and Konradsen 2017). Recent research has indicated that the intentional blending of V-labs with physical experiences has the potential for being especially effective in this regard (Olympiou and Zacharia 2012). In addition, if the assumption of role-taking proves to be accurate and V-Labs serve to make the practice of research more explicit and accessible, then such experiences could serve as a vehicle for supporting student persistence and success by aiding the development of their identity and confidence in relation to the professions of science and engineering (Chemers et al. 2011).

Since all of the necessary laboratory equipment is virtual, V-Labs also hold promise for institutions that are under-resourced for traditional laboratory facilities (Ogbuanya and Onele 2018). V-Labs can be completed in any physical space, such as a traditional classroom, lecture hall, or conference room (de Jong et al. 2013). Using V-Labs to offer laboratory courses in such spaces means that they can be completed by groups of students in arrangements that include the support of an instructor or teaching assistant without the need for costly, specialized and often difficult to schedule facilities (Tatli and Ayas 2013). V-Labs would also have utility in similar ways for online courses as long as students are provided with adequate social experiences and support. Thus, V-Labs can be

Fig. 1 Screenshot image of a VR Lab showing the student view as they work to determine blood type for a sample (image courtesy of Labster ApS)



viewed as representing a contemporary step in improving the accessibility of science and engineering education by minimizing the need for specialized laboratory infrastructure while maintaining the valued essence of the experience.

Over the past decade, immersive technologies have become much more affordable and feasible for educational applications (Martín-Gutiérrez et al. 2017). Though the potential of V-Labs has been extolled and investigated in a number of areas (e.g. Jones 2018), this research has not been synthesized for the specific context of undergraduate science and engineering education. For example, recent reviews of the literature pertaining more generally to virtual technologies have been completed for such topics as: (a) the general application in the context of K-20 education (Mikropoulous and Natsis 2011), (b) trends in available laboratory environments and their capabilities (Potkonjak et al. 2016), (c) training environments in the context of building evacuation (Feng et al. 2018), and learning outcomes of non-traditional laboratories compared to physical laboratories in the full K-20 context (Brinson 2015; Brinson 2017).

Our effort builds upon the work by Brinson (2015, 2017) who used a broad, K-20 perspective on participants and equally diverse collection of laboratory types in synthesizing studies from 2005 to 2015. While quite helpful as a synthesis of general learning outcomes due to their expansive collection of studies and KIPPAS analytical framework (i.e., Knowledge & Understanding, Inquiry Skills, Practical Skills, Perception, Analytical Skills, Social & Scientific Communication), the studies lack disaggregation by educational level, context, or specific nature of the experience, which are significant limiting factors. For example, V-Labs (as defined here) were classified together with other very different kinds of technology experiences such as simulations (see the methodology for how and why simulations were distinguished from V-labs) as well as those labeled as remote laboratories (see Ma and Nickerson 2006 for a discussion of the differences). The context and purpose of an undergraduate college or university course are significantly different from that of a typical middle or high school and combining these results is problematic. In addition, aside from aggregating findings for general learning outcomes, a synthesis of the perspectives, goals, contextual characteristics, and activities that were used in these studies was not provided. Therefore, a more nuanced approach focused on elements such as these was intended to illuminate the scope and limitations of the how, what, and why of our understanding about using V-Labs in this specific context.

Considering the global effort in higher education to adopt inquiry-based learning practices in laboratory education (Healey and Jenkins 2009; Mavromanolakis et al. 2014), as well as the increasing interest in innovative ways to support and engage students in the domains of science and engineering (see Holden and Lander 2012), a synthesis of V-Lab interventions in this context merits consideration. Thus, the goal

of this research was to focus only on the use of V-Labs in the specific context of undergraduate science or engineering education and to do so through systematic review and synthesis of existing peer-reviewed and published empirical studies. Further, in an effort to provide continuity across time and to support the differentiation of results based upon the nature of studies synthesized, we continue the use of Brinson's (2015) KIPPAS framework as one component of our analysis.

Due to the range of possible applications and perspectives that encompass undergraduate science and engineering education coupled with the rapid changes in technology that influence the form and function of what might have been defined as a V-Lab, we focused our study on the contextual characteristics that have been studied and the goals and perspectives that have influenced and emerged from this genre of empirical research during the time period 2009–2019. Contextual characteristics included the types of activities and outcomes that were investigated while the goals, perspectives, and interpretations were those explicitly indicated by the researchers in their published research papers. Accordingly, the following research questions framed our synthesis:

- 1 Which contextual characteristics, types of activities, and outcomes have been explored in studies of V-Labs?
- 2 What themes define the goals, perspectives, and interpretations used in this genre of research?

Methodology

This study involved a systematic review and synthesis of peer-reviewed empirical research papers published between 2009 and 2019 that were first identified through a search of seven relevant databases, then selected for inclusion based upon a defined criterion (Petticrew and Roberts 2006). Our approach employed elements of the protocol set forth by Khan et al. (2003) and according to our research questions, consisted of identifying relevant work, assessing the quality of studies, summarizing the evidence, and interpreting the findings. Each of these elements are detailed in the subsequent sections.

Identifying Relevant Work

The studies for review were located through systematic searches of ProQuest, Educational Journal, Wiley Online, APA PsyNet, Web of Science, JSTOR, and Applied Science and Technology databases where only peer-reviewed empirical journal articles were deemed eligible. The use of published journal articles as acceptable evidence for review was based upon the assumed level of quality that is provided by the peer-review process. As well, journal articles tend to provide a more complete description and analysis of data, as opposed

to conference proceedings that may involve only the early stages or preliminary forms of each.

The search process was conducted using the following terms: “virtual reality” and undergraduate and “STEM,” “virtual learning” and undergraduate; “virtual laborator*” and undergraduate; “simulation learning”; “virtual simulation” and undergraduate; “computer simulation” and undergraduate and science; “computer simulation” and undergraduate and engineering. To limit results to the most up-to-date research, additional parameters included limiting the applicable date range to January 2009 to December 2019. This process resulted in 1653 potential articles which were then subjected to our additional criteria in two additional steps based upon a review of abstracts and then screening the full articles.

Inclusion-Exclusion Criteria

The inclusion and exclusion criteria were constructed based upon the purpose and research questions, which were then modified slightly as the abstracts were read and issues and opportunities were discovered. As part of our exploratory searches to identify relevant studies, we recognized the need for a primary criterion based upon a clear distinction between V-Labs and what are more commonly described as simulations. Since this directly involved our operational definition of V-Labs, we begin here by explaining the nature of this distinction and why studies of simulations were excluded from our corpus of research as the first stage in our process.

The term simulation and some form or iteration of the phrase virtual laboratory are oftentimes used interchangeably and synonymously in reports of research when referencing certain educational technologies. For instance, Ma and Nickerson (2006) defined virtual laboratories as, “imitations of real experiments...[where]...the infrastructure required for laboratories is not real but simulated on computers” (pg. 6). This definition was then used by Brinson (2015) who’s synthesis included studies that used both simulation and virtual laboratories but were couched under the term virtual laboratory because of the fully online nature of the experience. Though the terms share aspects, there is a clear distinction in terms of what they describe. Both are based upon computational models, but simulations typically do not involve solving a scientific problem, but are constrained in such a way so as to focus users on manipulating and/or modifying a given set of parameters as a precursor to understanding the effect on other variables in the form of an output. Examples of simulations include the many options available from the PhET project at the University of Colorado (<https://phet.colorado.edu/>). An exception to this would be simulations in nursing or medical education where the computational model can take the form of a virtual patient or a haptically enabled mannequin which are lifelike, but not situated in an emulation of the real-world (Rush et al. 2010).

In medical training programs, the simulation is used more broadly to define a purposefully constructed situation involving a role-player/actor, virtual patient, or haptically enabled mannequin, such that the term represents a portrayal of a relevant actual event or situation (Rush et al. 2010). An example of this would be Shadow Health (www.shadowhealth.com), which creates virtual patients that are designed to train medical staff on interviewing and examining patients. Such a simulation allows users to use natural language to engage with the virtual patients and receive answers based on the questions asked. In domains such as architecture, engineering, and construction, simulation refers to creating prediction models with the intent of determining accurate outcomes that can then be used to inform decision-making (Akhavian and Behzadan 2015).

Studies that involved simulations were not included in our review and applying this criterion excluded 970 manuscripts. The remaining 88 manuscripts were subjected to the following additional criteria. Studies that reported the results of an empirical study involving undergraduates as participants using a V-Lab in either 2D or 3D format as part of laboratory learning and were published between January 2008 and December 2018 were included and those that did not meet this criterion were excluded (e.g., Schott and Marshall 2018; Uribe et al. 2016). Manuscripts that did not include the collection and analysis of data on student outcomes (e.g., learning, perception, self-efficacy, motivation) were excluded (e.g., Dalgarno et al. 2009; Fang and Tajvidi 2018; Kang et al. 2018; Williams 2010). Additionally, Achuthan et al. (2017); Reyes-Aviles and Aviles-Cruz (2018); and Tetour et al. (2011) were excluded because these studies focused on remote experiments and/or augmented reality environments. After applying the full complement of inclusion and exclusion criteria, the corpus of data consisted of 25 articles, which were published in 23 different journals that cover the subjects of science and engineering. These articles were then subjected to our content analysis procedure (Fig. 2).

Analysis Procedure

The initial phase of content analysis consisted of a coding process for study characteristics based upon our research questions where individual narrative or representational elements of each manuscript were selected as evidence (Krippendorff 2012). Elements of the manuscripts that addressed our research questions were highlighted and given a name (code) that represented a characteristic. Characteristics and example codes included the: general domain of science or engineering (e.g., chemistry, biology); nature and form of technology used (e.g., 2D, 3D); specific domain or course context (e.g., microbiology, manufacturing, introductory physics for engineers); study design (e.g., stated purpose or goal, definition of virtual reality, research questions,

constructs investigated); theoretical framework and/or major areas of research reviewed for framing the study (e.g., motivation, cognitive theory of multimedia learning); methodology (e.g., qualitative, quantitative, mixed); forms of data (e.g., interviews, surveys); results (e.g., charts, graphs, knowledge claims). In addition, student outcomes were coded according to the KIPPAS framework (Brinson 2015), which consists of six categories for the types of student outcomes: Knowledge & Understanding (e.g., Quiz/Exam), Inquiry Skills (e.g., testing a hypothesis), Practical Skills (e.g., Lab Practical), Perception (e.g. Survey/Questionnaire), Analytical Skills (e.g., performing an analysis), and Social & Scientific Communication (e.g., Lab Report/ Written Assignment). As an analytical tool, the KIPPAS was designed to reflect the goals and laboratory experiences identified by the National Research Council in two separate reports (NRC 2006, 2012). The KIPPAS was also designed to capture the frequency of outcomes assessed, therefore, if one study assessed two outcomes, both would be represented. The identified characteristics were thus used as a means for making the epistemic elements of each study explicit and accessible for comparison.

The authors met regularly to code, categorize, and discuss emergent patterns and themes to consensus (Lincoln and Guba 1985) and a synthesis table was constructed to organize the characteristics and improve our interpretation of the commonalities among the studies (Table 1). We began our synthesis by building an understanding of the why and what concerning the use of V-Labs, then proceeded to the how. Exploring the why and what resulted in our theme of *Perspectives on Implementing V-Lab Experiences*, which involved the expectations that were used as a means for predicting or hypothesizing the impact of V-Labs. We inferred that how someone planned on using something was most likely influenced by their perception of its usefulness or intent for how it should be used. Exploring the how of these studies resulted in our theme of Interpretation of V-Lab Experiences.

In the following sections, we present our results based upon the research questions. We begin by defining the content domains, contexts, and outcomes, which is followed by a synthesis of the perspectives, goals, and interpretations.

Results

The majority of studies fell within the general domain of science (60%) and to a lesser extent engineering (40%), which is consistent with prior results from Mikropoulos and Natsis (2011) (Table 2). Ninety-two percent of the learning environments described in these studies were of the 2D desktop variety (e.g., Hawkins and Phelps 2013) and 75% were acquired from an outside source/vendor and were not exclusively designed for the course in which they were implemented. The

vendor for the V-Labs was also not exclusive, with only seven articles sharing the same vendor (i.e., Labster), but for different V-Lab experiences. While the selection criteria required articles to be situated in undergraduate education, 26% reported involving an introductory course (e.g., Olympiou and Zacharia 2012) and for 80% of the total, this was a single experience that was related to a module or activity within the course.

As table one suggests, the overall evaluation focus in lieu of research is dominant. In these cases, the authors identified the V-Lab condition as playing a role in the learning process and subsequently measured the impact on some form of student outcome. The studies typically lacked a theoretical perspective or in some cases even research questions that could have been used for interpreting changes in student outcomes. A positive change in content knowledge was the most often targeted student outcome (Fig. 3). For instance, Chini et al. (2012) explored how the use of a V-Lab would help students understand science concepts related to pulleys while Darrah et al. (2014) evaluated students' content knowledge after experiencing a V-Lab. Exploring student perceptions of V-Labs was also a common goal for these studies. For example, Dyrberg et al. (2017) studied science students' attitude with regard to their use of a V-Lab while Koh et al. (2010) explored how the use of a V-Lab in an engineering course would influence students' motivation to learn the course content.

Although improving practical or inquiry skills is considered an affordance of laboratory education, it was not widely studied (24%) with V-labs. Examples of when it was, included the process of analyzing protein expression by muscle cells (Polly et al. 2014) or practicing how to properly generate x-ray images with virtual patients (Gunn et al. 2018). This finding is also consistent with those of Brinson (2015).

Perspectives on Implementing V-Lab Experiences

More than half of the studies (60%) were based upon an expectation that V-Labs would bring about some predetermined student outcome of interest. A number of studies based this expectation on V-Labs functioning as inquiry activities, assuming that students would find this interesting and would take advantage of opportunities for practice with laboratory methods. For instance, Polly et al. (2014) predicted that a V-Lab would improve students' ability to apply molecular laboratory techniques as well as to repeat and redesign experiments. Cheong and Koh (2018) indicated that a V-Lab could be used by students to solve engineering related math problems and Ogbuanya and Onele (2018) hypothesized that a V-Lab would influence deeper learning through engineering practice.

No consistent pattern was identified for the remaining studies, as they were driven by a variety of perspectives. For instance, Chini et al. (2012) was based upon the expectations of

Table 1 Overview of papers used for the review

Citation	Domain/course	Framework (theoretical/conceptual)	Research focus	Constructs	Results
			Eval Exp		
Ángel (2015)	Chemical engineer	None stated	●	Learning gains, motivation	Students had a better scientific and technical understanding of bioreactors. V-Lab increased student learning compared to traditional laboratory
Bortnik et al. (2017)	Analytical chemistry	Cognitive development & experiential learning and meaningful learning	●	Scientific literacy and research skills	Results supported the use of V-Lab in conjunction with a physical laboratory as beneficial in enhancing student research skills and practices in analytical chemistry
Cheong and Koh (2018)	2nd year engineering math	Visible thinking	●	Perception of effectiveness and overall reception	V-Lab was perceived to be greatly beneficial at enabling students to apply learned knowledge to authentic problems
Chini et al. (2012)	Intro Physics for non-science majors	Dynamic transfer	●	Understanding of concept, benefit of modality in relation to content learned/timing of post-test/sequencing	Either physical or virtual laboratory will be more appropriate for a given experiment depending on the concept being taught
Cobb et al. (2009)	Bioscience	None stated	●	Student perception, perceived impact on learning	V-Lab does not reduce the learning gain following teaching by traditional methods
Darrah et al. (2014)	Intro physics	None stated	●	Content knowledge	V-Lab produced similar learning outcomes to a hands-on traditional laboratory
Dyrberg et al. (2017)	Pharmaceutical toxicology and microbiology	Motivation	●	Task value, self- efficacy, and content knowledge	Students' perceived value of engaging in the V-Lab is lower than the perceived value of engaging in a real laboratory
Ekmekci and Gulacar (2015)	Secondary STEM	Computer-based instruction	●	Learning gains, instruction characteristics	The use of hands-on laboratory compared to a V-Lab produced similar results for learning gains
Gao et al. (2011)	Hydraulic engineering	None stated	●	Completion time, student perception	Students using a V-Lab completed physical experiments in a shorter time. Students perceived the V-Lab to positively influence their learning
Goudsouzian et al. (2018)	Cell and molecular biology	Design thinking and agile	●	Learning gains, student perception, self-efficacy	Learning gains between the two conditions showed similar growth. Students have an improved sense of self- efficacy in response to performing the physical laboratory and V-Lab
Gunn et al. (2018)	Radiographic practice	None stated	●	Technical skill, role play	Students using a V-Lab were not at a disadvantage compared to those using a traditional medical simulation
Hawkins and Phelps (2013)	General chemistry	None stated	●	Conceptual understanding and transfer	V-Lab was just as effective as a physical laboratory at teaching general chemistry concepts
Koh et al. (2010)	Mechanical engineering	Self-determination	●	Motivation, learning orientation, student performance	Students in both V-lab and physical laboratory conditions perceived their psychological needs to be satisfied
Makransky et al. (2019)	Biology	Equivalence hypothesis	●	learning outcome, intrinsic motivation, self-efficacy	Students learning from a V-Lab independently while at home performed just as well as those who used a V-Lab in class
Makransky and Lilleholt (2018)	Biology	Control value theory of achievement emotions	●	Presence, motivation	Students reported an increased preference for V-Lab with a head mounted display versus a V-Lab with a desktop computer
Makransky et al. (2017)	Not stated	CTML (redundancy)	●	Learning gains, brain activity	Students' experienced a greater sense of presence in the high immersion but learned less than students who experienced less-immersion. There was no difference between text only compared to text and narration
Makransky et al.	Microbiology	None stated	●	Content knowledge, intrinsic motivation, self- efficacy, lab scores	V-Lab was successfully in preparing students for a physical laboratory

Table 1 (continued)

Citation	Domain/course	Framework (theoretical/ conceptual)	Research focus		Results
			Eval	Exp	
(2016) Nolen and Koretsky (2018)	Engineering	Scientific teaching (active learning, student assessment and inclusive teaching)	•	Self-reported interest, student perception of engagement	Students using a V-Lab reported greater engagement and perception of contributions to their group's learning. Those who used the V-Lab displayed a greater positive impact on end-of-course interest in engineering
Ogbuany and Onele (2018)	Electrical engineering	Inquiry-based teaching	•	Academic achievement, learning interest, learning engagement	Students who experienced the V-Lab had a significant difference in academic achievement, learning interest, and learning engagement compared to the traditional laboratory
Olympiou and Zacharia (2012)	Physics by inquiry	PM VM	•	Student understanding	V-lab and physical laboratory together is more conducive to laboratory learning than either alone
Polly et al. (2014)	Not stated	Inquiry-based learning	•	Student understanding, diagnostic skills, learning of research skills	V-Lab improved student perceptions of certain aspects of their learning. However, was equivalent to the physical laboratory in all aspects evaluated
Secomb et al. (2012)	Nursing skills laboratory	Cognitive constructivism	•	Content knowledge, cognitive skills	Activities completed in a V-Lab did not develop cognitive skills of undergraduate nursing students compared to traditional methods
Toth (2016)	Cellular machinery course	None stated	•	New knowledge construction, student learning	Students were able to develop and apply new knowledge with the V-Lab. However, students continued to struggle with reasoning
Vrellis et al. (2016)	Not conducted in a course	Problem-based learning	•	Learning outcome and satisfaction	Both the V-Lab and traditional laboratory evoke the same overall learning outcome and satisfaction
Zhu et al. (2018)	Offshore oil & gas engineering	None stated	•	Practical operation, content knowledge, and perception	Students using V-Labs displayed increased practical and theoretical knowledge. Students perceived the V-Lab to enhance their understanding and practical application

Note: Eval = evaluation study; Exp = experimental study

dynamic transfer—the capacity to use new knowledge and problem-solving skills in a unique context—they endeavoring to understand how the use of support elements within the V-Lab such as alternative interpretations and feedback influenced students' problem-solving skills in physics. In an effort to provide “alternative interpretations and feedback” (Chini et al. 2012 p.3), the V-Lab provided immediate feedback to responses concerning abstract concepts. Similarly, Vrellis et al. (2016) inquired how a problem-based activity situated in a V-Lab would impact students' cognitive and non-cognitive outcomes. Ideally, the process that students experienced solving an authentic content related problem in the V-Lab would result in improved knowledge construction. Likewise, Secomb et al.

(2012) explored the influence of a V-Lab on students' scientific reasoning which entailed explaining and analyzing results, effect change, and scientific arguments. It was hypothesized that the engaging and interactive environment presented in the V-Lab would promote “effective learning” (p. 3476).

Goals for Implementing V-Labs

The role and significance of laboratory learning have evolved over the last century and the goals for implementing V-Labs have emerged as a product of this process. As our understanding of how students learn science has expanded, so has the emphasis on laboratory education, the “experiences [that]

Table 2 A breakdown of studies by subject within the general domain of science and engineering

Number	Discipline	Citation
4	Biology	Dyrberg et al. (2017), Makransky and Lilleholt (2018), Makransky et al. (2019)
3	Biotechnology	Cobb et al. (2009)
3	Cellular and microbiology	Goudsouzian et al. (2018); Makransky et al. (2016); Polly et al. (2014); Toth (2016)
2	Chemical engineering	Ángel (2015); Nolen, and Koretsky (2018)
2	Chemistry	Bortnik et al. (2017); Hawkins and Phelps (2013)
2	Civil engineering	Cheong and Koh (2018)
2	Clinical nursing practicum	Secomb et al. (2012)
1	Electrical engineering	Ekmekci and Gulacar (2015); Ogbuanya and Onele (2018)
1	Hydraulic engineering	Gao et al. (2011)
1	Mechanical engineering	Koh et al. (2010)
1	Not stated or not within a course	Makransky et al. (2017); Vrellis et al. (2016)
1	Offshore oil & gas engineering	Zhu et al. (2018)
1	Physics	Chini et al. (2012); Darrah et al. (2014); Hawkins and Phelps (2013)
1	Radiographic	Gunn et al. (2018)

provide opportunities for students to interact directly with the material world (or with data drawn from the material world), using the tools, data collection, techniques, models, and theories of science” (Singer et al. 2005 p. 31). From this corpus of research, our analysis indicates that researchers translate the potential promise of V-Labs into two primary goals for their work. First, V-labs are implemented with the intent of serving as an approach to teaching, which differs from them as a resource or element of curriculum. Second, V-Labs are

intended as a vehicle or intervention for improving student outcomes.

As V-Labs have become more accessible to the general public, principally due to the reduction in cost for the technology, researchers have increasingly focused on examining their viability in what we interpret as a teaching approach; a means to “pass acquired knowledge, skills, and technology between individuals” (Fogarty et al. 2011). A number of studies proposed that V-Labs were an improvement to how content and/or skills had previously been taught where the V-Lab was

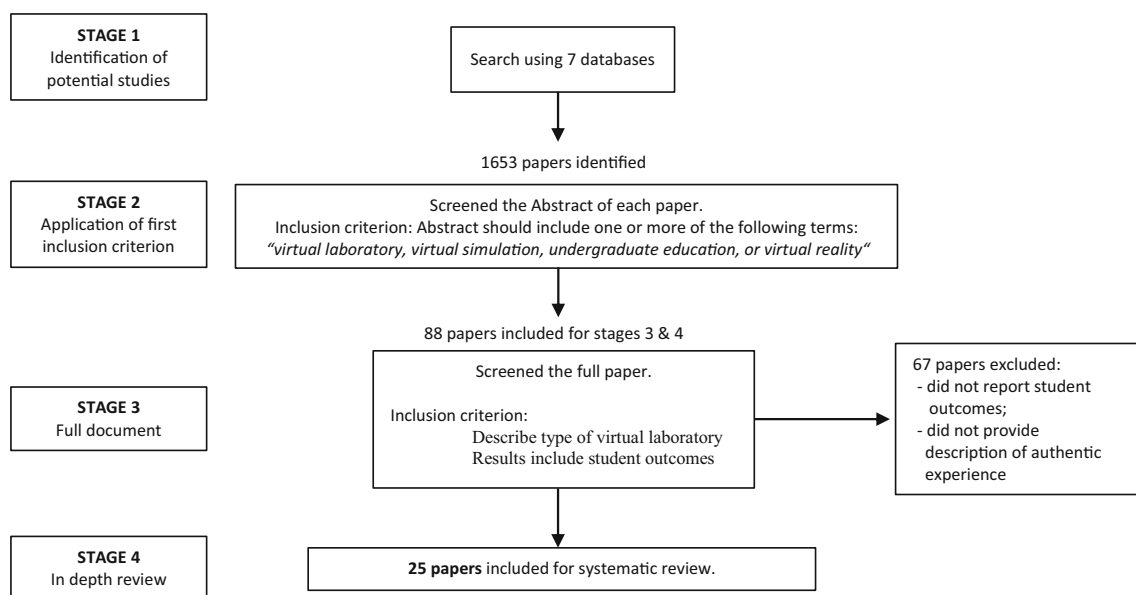


Fig. 2 Inclusion and exclusion criteria and their impact on the number of studies reviewed

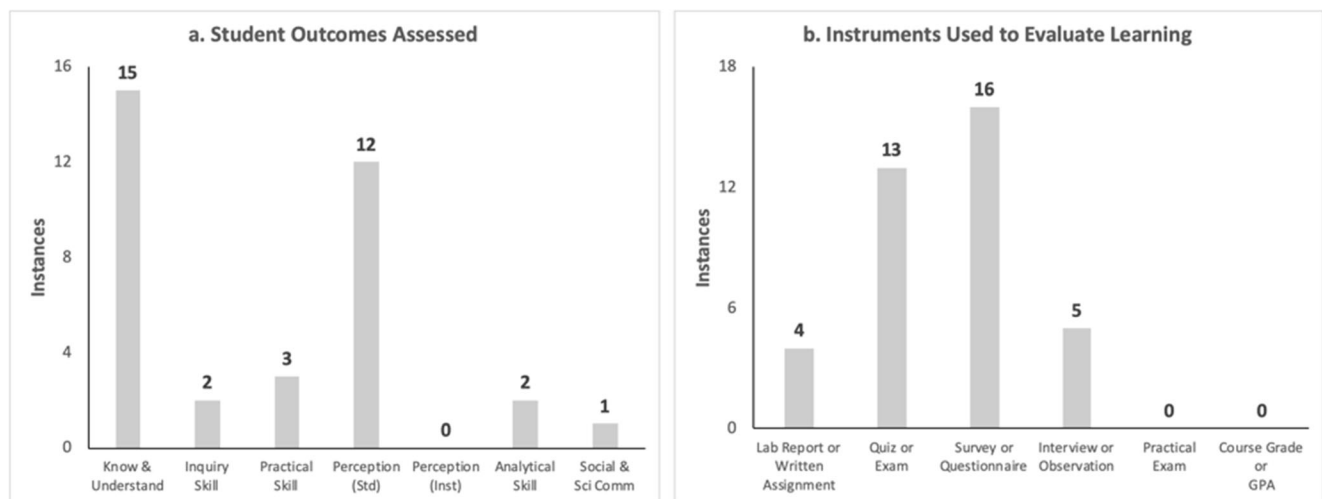


Fig. 3 The **a** types of student outcomes that were assessed and **b** instruments that were used to evaluate learning across all 25 studies. Some studies assessed numerous outcomes with multiple instruments

assumed to be more beneficial than traditional instructional methods (Ekmekci and Gulacar 2015). In these instances, the student-focused nature of V-Labs is promoted as a way to improve the learning of the science or engineering content (Goudsouzian et al. 2018; Ekmekci and Gulacar 2015). As a teaching approach, V-Labs are indicated as a means for increasing student satisfaction, despite situations of limited physical resources (Cobb et al. 2009) or instances where the V-labs expand on the locations where laboratory learning can take place (Bortnik et al. 2017).

Ángel (2015) studied the capacity of V-Labs for providing instructors with an opportunity to eliminate language barriers that can occur in laboratory learning for students who are not native English speakers. In an effort to move away from passive teaching methods, those that are based upon merely transferring knowledge from the instructor to the students, V-Labs have been investigated as an active teaching alternative (Michel et al. 2009). For example, Cheong and Koh (2018) sought to understand how a V-Lab could improve “general passivity” (p. 58321) in a classroom by investigating whether the V-Lab’s replicated real-world scenarios would support students’ having thoughtful discourse and engagement, which would result in a greater capacity to apply their knowledge. Similarly, Goudsouzian et al. (2018) and Toth (2016) acknowledged the need for multiple teaching approaches when disseminating information to students and proposed that the affordances of V-Labs (i.e., animation, videos, and teaching aids) were well suited for improving learning outcomes. “[V-Labs] are software-tools that allow users to design repeated experiments to test the effects of variables... but in a shorter amount of time, with increased safety, and at a reduced cost” (Toth 2016 p. 158).

V-Labs were promoted as being poised to provide different instructional formats, especially in the case of blended

learning. In the study by Bortnik et al. (2017), students used a V-Lab with traditional in-class instruction as a supplement prior to completing their hands-on laboratory. Researchers identified perceived affordances of the V-Lab (i.e., student-centered, inquiry-based, affordable, and expanded access) as being beneficial to improving students “scientific literacy, research skills, and practices” (Bortnik et al. 2017 p.3). In a similar fashion, Goudsouzian et al. (2018) sought to explore an expanded teaching approach using V-Lab as a way to disseminate course content to students in a more engaging manner. Citing a limited availability of laboratory activities for illustrating microbiology topics such as the cell cycle, a V-Lab was determined to be ideal, providing learning opportunities where none previously existed (Goudsouzian et al. 2018). “Multiple pedagogical approaches are useful in helping undergraduate students learn difficult scientific concepts, and incorporating multisensory learning tools aids in the ability of individuals to recall information” (Goudsouzian et al. 2018 p. 361).

A number of studies present V-Labs as being wholly responsible for providing instruction, akin to being the teacher. However, other studies acknowledge that learning using a V-Lab should only be seen as an instructional tool, not supplanting the teacher, but something for a teacher to use (Ekmekci and Gulacar 2015; Vrellis et al. 2016). Yet, many of the actions and activities that are typically provided by a teacher or laboratory instructor, such as feedback during work sessions are being subsumed by features of V-Labs and lauded as an advantage (Cobb et al. 2009; Bortnik et al. 2017). For example, V-Labs are tasked with providing content knowledge (Ogbuanya and Onele 2018; Bortnik et al. 2017) with the intent of assisting students in memorizing information (Goudsouzian et al. 2018) or providing students with knowledge of “how to streak out bacteria on agar plates to isolate

single colonies [and] described the principles of using selective and differential culture media” (Makransky et al. 2016 p.11). To assure the fidelity of a V-Lab for providing instruction similar to a teacher, it is critical to examine the components of how students engage with their instructor(s) during a lesson and exploring to what extent the implementation of the V-Lab supports the identified engagement components.

As an intervention intended to improve student outcomes, studies were categorized into one of two different themes based upon the overall intent for using a V-lab. Either the V-Lab was used as a replacement to an existing learning activity in a traditional laboratory or as a supplement. The theme of replacement included a few studies that also explored the sequencing of V-Labs with physical laboratories. In sum, each of these goals resulted in mixed outcomes (Table 3).

The theme of replacement involved the use of V-Labs as an assumed equivalent or greater learning experience and for these studies, the most likely finding was no significant difference from a traditional laboratory. This was the case in the context of a microbiology course with content knowledge and self-efficacy as outcomes (Makransky et al. 2016), in the context of electrochemistry assessing conceptual understanding (Hawkins and Phelps 2013), and with physics (Darrah et al. 2014) and electrical circuits more specifically (Ekmekci and Gulacar 2015). However, two replacement studies involving an electrical circuit laboratory (Ogbuanya and Onele 2018) and offshore practical training (Zhu et al. 2018) did report an improvement in content knowledge as well as increased interest in favor of V-Labs. In a few cases, the notion of replacement was extended and modified into an approach that involved the sequencing of V-Labs and alternating them with traditional forms of laboratory. For example, Olympiou and Zacharia (2012) found that students who experienced a blending of physical laboratories with V-Labs indicated enhanced

understanding of the content. Chini et al. (2012) employed a similar approach of concept and implementation sequencing but did not find a positive influence on student outcomes.

The theme of supplement refers to an investigation into whether adding a V-Lab was beneficial as an additional learning experience without taking anything away. When used in this fashion, a positive influence on student outcomes was more likely found than not. For example, Dyrberg et al. (2017) reported an increase in feelings of confidence and comfort when operating the physical laboratory equipment after use of a V-Lab supplement, while Bortnik et al. (2017) documented enhanced research skills and practices. In the case of Nolen and Koretsky (2018) the researchers inquired if the perceived affordances of a V-Lab such as, the ability to provide, “wide range of length scales, time scales, and complexity” (p. 227) would influence student engagement in an undergraduate engineering course. Results indicated an increase in interest and engagement in the course content when using the V-Lab. Similarly, Goudsouzian et al. (2018) found that students who completed either a live or V-Lab also made gains in laboratory skills and experimental predictions while Toth (2016) found an increase in students’ knowledge scores after using a V-Lab, but not when asked to apply information.

Interpretations of V-Lab Experiences

Aside from evaluating student outcomes of interest, two larger frameworks were noted for interpreting the influence of V-Labs, motivation (generally, as well as expectancy value, self-determination theory, and control value theory of achievement emotions specifically) and the Cognitive Theory of Multimedia Learning (CTML). The theories used to interpret the impact on student outcomes suggests that researchers inferred the primary impact as either being

Table 3 A breakdown of effects on student outcomes by overall goal of the study

Goal of the study	Studies indicating a positive effect on student outcomes	Studies indicating no effect or negative effect on student outcomes
Replacing a traditional laboratory experience	Ángel (2015); Cheong and Koh (2018); Cobb et al. (2009); Gao et al. (2011); Makransky et al. (2018) Nolen and Koretsky (2018); Ogbuanya and Onele (2018); Toth (2016); Zhu et al. (2018)	Darrah et al. (2014); Dyrberg et al. (2017); Ekmekci and Gulacar (2015); Gunn et al. (2018); Hawkins and Phelps (2013); Koh et al. (2010); Makransky et al. (2017); Makransky et al. (2019); Polly et al. (2014); Secomb et al. (2012); Vrellis et al. (2016)
Supplementing a traditional laboratory experience	Bortnik et al. (2017); Makransky et al. (2016)	Goudsouzian et al. (2018)
Sequencing VR Labs with traditional laboratory experiences	Chini et al. (2012); Hawkins and Phelps (2013)	Ekmekci and Gulacar (2015)

personally motivating for students or for improving their capacity for processing information. A modest number (15%) of articles in the review apply theories of motivation for interpreting undergraduate student experiences with V-Labs pursuing to extend understanding to include non-cognitive factors. For example, Dyrberg et al. (2017) acknowledged students' perception of a task influenced their level of engagement and explored if they found value in the use of V-Lab as a pre-laboratory exercise. Koh et al. (2010) documented the extent to which a V-Lab influenced participants' psychological needs and learning outcomes. Nolen and Koretsky (2018) explored the influence a V-Lab would have on student engagement in an engineering course. The V-Lab experience was designed to improve participants' ability to work in a team and collectively engage in an intricate task, as well as increase their interest in the subject matter. Likewise, Makransky and Lilleholt (2018) investigated how an individual's perceived level of immersion influenced their "non-cognitive and perceived learning outcomes" (p. 1144).

The utilization of process related theories occurred in a single study that applied the Cognitive Theory of Multimedia Learning (CTML), which contends that a selective amount of both visual and verbal representation is ideal for a conducive learning experience (Mayer 2014). In Makransky et al. (2017), it was hypothesized that learning outcomes would be a direct reflection of how increased immersion influenced students' extraneous processing and that V-Labs that include both narration and written text would negatively influence student outcomes. The results supported this prediction, showing higher cognitive load and lower student learning in the V-lab condition.

Discussion

This review reveals a dearth of varied theoretical and methodological approaches regarding V-Labs. Nearly half of the articles did not explicitly state a theoretical perspective in interpreting student outcomes or offered research questions that guided the study. Lacking a theoretical perspective eliminates the critical assessment of how a pedagogical practice such as implementing a V-Lab could influence forms of engagement or knowledge construction. In essence eliminating the capacity to associate any change in performance with a specific process for explaining why the change occurred (Driscoll 2005). Additionally, the reviewed articles lacked the use of a consistent definition for V-Labs, one that (a) might have accounted for the intended experience, (b) included established characteristics and critical features, or (c) was not just a description of the technology used.

The majority of the studies were evaluative with most seeking to establish the efficacy of V-Labs as an option for meeting the specific needs of a particular context. Pre-and post-tests in

the forms of a survey, quiz or exam were used to measure if the V-Lab, serving as an intervention, increased content knowledge when compared to another group. This form of baseline assessment is customary in technology integration, but does not provide a rich understanding of the teaching or learning context in ways that can advance our capacity to explain and account for any differences that might be detected. The current corpus is devoid of studies that take more than a minimal cognitive approach to understanding student learning. All of these studies were completed from the perspective that a V-Lab alone would serve as a learning intervention by replacing all human interaction (e.g., peer-to-peer, peer-to-instructor). This suggests a need for future studies that take a multidimensional construct- or person-oriented approach, one that recognizes learning as more than acquired content knowledge and explores the importance of social interactions in relation to V-labs. For example, teaching assistants are a mainstay of traditional undergraduate teaching laboratories and prior research indicates their importance (Gardner and Jones 2011), yet it seems that most V-labs are designed and subsequently investigated as if these individuals are not important or simply do not exist. Teamwork and collaboration among students is widely recognized as a key element of laboratory learning (Bauerle et al. 2011; Committee 2018; Hofstein 2004), but is not addressed as part of how V-labs are designed or how they are applied or researched. The teaching laboratories that V-Labs are intended to emulate include people who mentor, coach, collaborate, counsel, and instruct. These elements of the broader learning environment need to be better addressed in both the design of V-labs as well as research related to their use.

A small number of select quantitative studies did go beyond basic efficacy and should serve as models for future studies. Makransky et al. (2017, 2020); Makransky and Lilleholt (2018) showed promising results in their exploration of the influence of V-Labs on non-cognitive outcomes such as motivation and cognitive load. Yet, we found no attempts at replicating or building upon these studies. Studies such as these should serve as a starting point and initial model for how quantitative studies should be conducted in order to further develop our understanding of the phenomenon. New replication studies would validate the reported findings and confirm the results as broadly applicable. Future research should emphasize building from these methodologies to include multiple iterations, various V-Lab vendors, diverse domains, and varied forms of data. Quantitative research needs to move beyond the paradigm of only evaluating acquisition of content knowledge and skills with methods that involve comparison of V-lab with real-world experiences under the assumptions of equivalence.

The first-person perspective of participants, including students, teaching assistants and instructors is noticeably missing from the existing research on V-labs. Student perceptions were most often captured using survey instruments that

offered minimal opportunities for free response. This represents a lost opportunity, producing a situation that may result in reinforcing or validating problematic assumptions about learning, including the role of people in the process. Given the predominance of using off-the-shelf software, this would be particularly pertinent and implies that participant needs and expectations are not being assessed or subsequently addressed. There seems to be a universal assumption that students view V-labs as interesting and important, expect to be successful in navigating and learning from them, find inherent value in this type of task, and have confidence in what they know and are able to do. This assumption seems to be grounded more in presumptions about participants being motivated by the novelty of the technology than the learning attributes that would be afforded by the design of any learning environment (Wells et al. 2010). In their description of using a fidelity of implementation process for constructing evidence-based practices, Stains and Vickrey (2017) acknowledge that simply inserting a new approach such as V-lab and measuring the student outcomes ignores vital elements of the implementation process that can influence a study's results.

Studies that detail and describe the variation in student experience with V-Labs are sorely needed. In particular, those that explore the influence of individual background or personal characteristics on the V-Lab experience and how design can be used to afford social learning and models of instruction that include peers and instructors. The meaning that students make of their experience with V-labs is largely undocumented and the design of the available environments is all based upon the assumption that students can discern the critical features when appropriate in order to derive the intended meaning. Existing software designs are heavily influenced by instructional perspectives, the intent to emulate the attributes and affordances of an existing physical space, and to provide a singular participant experience. Studies of this type would focus on using qualitative inquiry to document the variation in student experiences in order to provide key information about the features that students attend to as they work through a V-lab and then use that experience to make personal meaning (Bussey et al. 2013). Insights derived from these studies would offer the potential for new approaches to design, ones that meld the expectations of students with the goals of instructors and designers. Better accounting for the varied student perspectives in such a manner would result in collectively moving from “does [it] work”, to “why, how, and under what conditions could [it be] impactful” (Stains and Vickrey 2017 p. 2).

Limitations

Our search process was limited by our choice of search engines as well as our ability to identify and use search terms that were consistent with how researchers defined their work. Thus, it is possible that our search process did not identify

all of the published studies on V-labs during the time period. However, even in such a scenario, this review can still be viewed as a relative and likely representative sample of the available literature, which exposes a need for much more robust and diversified research regarding the use of virtual reality in undergraduate STEM laboratory education.

Conclusion

Use of V-Labs has shown some promise for expanding the capabilities of laboratory education. However, the results of this study reveal a dearth of theoretical and methodological approaches used for exploring this phenomenon. Most of the studies were evaluative, seeking to establish the efficacy of V-Labs for meeting specific needs under the expectation that simply using them would bring about predetermined student outcomes. V-Labs were used primarily as a teaching approach, responsible for providing instruction without any need or designed intent for interaction with a human teacher or peers. Regardless of approach, these elements are consistently shown to be key predictors of student learning. Interpretation of outcomes was most likely based upon an assumption that students would find them personally motivating. An assumption that seems more grounded in a novelty effect than design of the environment. New studies are needed that explore how individual background or personal characteristics influence the variation in V-Lab experience and how design can be used to afford social learning with current models of instruction through the inclusion of teachers and peers. Exploring such questions would offer a richer perspective of virtual laboratories in undergraduate education.

Acknowledgments The authors wish to thank Dr. Leonard Annetta for his service as Guest Editor for this manuscript as well as the helpful comments from the anonymous reviewers.

Compliance with Ethical Standards

The research reported here involved analysis of existing published studies as primary sources and thus did not involve human participants.

Conflict of Interest Kent J. Crippen serves as the Editor-in-Chief of this journal. To mitigate this conflict of interest, Leonard Annetta, a senior member of the Editorial Board assumed the role of Guest Editor for this manuscript and was responsible for all decision making.

References

- Achuthan, K., Francis, S. P., & Diwakar, S. (2017). Augmented reflective learning and knowledge retention perceived among students in classrooms involving virtual laboratories. *Education and Information Technologies*, 22(6), 2825–2855. <https://doi.org/10.1007/s10639-017-9626-x>.

- Akhavian, R., & Behzadan, A. H. (2015). Construction activity recognition for simulation input modeling using machine learning classifiers. In *Proceedings - Winter Simulation Conference* (Vol. 2015-January, pp. 3296–3307). Institute of Electrical and Electronics Engineers Inc.. <https://doi.org/10.1109/WSC.2014.7020164>.
- Alexander, P. (2003). The development of expertise: the journey from acclimation to proficiency. *Educational Research*, 32(8), 10–14.
- Ángel, S. A. (2015). Real and virtual bioreactor laboratory sessions by STSE-CLIL WebQuest. *Education for Chemical Engineers*, 13, 1–8. <https://doi.org/10.1016/j.ece.2015.06.004>.
- Baker, R. S., Clarke-Midura, J., & Ocumpaugh, J. (2016). Towards general models of effective science inquiry in virtual performance assessments. *Journal of Computer Assisted Learning*, 32(3), 267–280. <https://doi.org/10.1111/jcal.12128>.
- Bernhard, J. (2018). What matters for students' learning in the laboratory? Do not neglect the role of experimental equipment! *Instructional Science*, 46(6), 819–846. <https://doi.org/10.1007/s11251-018-9469-x>.
- Basey, J. M., Mendelow, T. N., & Ramos, C. N. (2000). Current trends of community college lab curricula in biology: an analysis of inquiry, technology, and content. *Journal of Biological Education*, 34(2), 80–86. <https://doi.org/10.1080/00219266.2000.9655690>.
- Bauerle, C., DePass, A., Lynn, D., O'Connor, C., Singer, S., Withers, M., et al. (2011). In C. A. Brewer & D. Smith (Eds.), *Vision and change in undergraduate biology education: a call to action*. American Association for the Advancement of Science (AAAS) Retrieved from <https://live-visionandchange.pantheonsite.io/reports/>.
- Bortnik, B., Stozhko, N., Pervukhina, I., Tchernysheva, A., & Belysheva, G. (2017). Effect of virtual analytical chemistry laboratory on enhancing student research skills and practices. *Research in Learning Technology*, 25. <https://doi.org/10.25304/rlt.v25.1968>.
- Brinson, J. R. (2015). Learning outcome achievement in non-traditional (virtual and remote) versus traditional (hands-on) laboratories: a review of the empirical research. *Computers and Education*, 87, 218–237. <https://doi.org/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. *Journal of Science Education and Technology*, 26(5), 546–560. <https://doi.org/10.1007/s10956-017-9699-8>.
- Bussey, T. J., Orgill, M., & Crippen, K. J. (2013). Variation theory: A theory of learning and a useful theoretical framework for chemical education research. *Chemistry Education Research and Practice*, 14(1), 9–22. <https://doi.org/10.1039/c2rp20145c>.
- Chemers, M. M., Zurbriggen, E. L., Syed, M., Goza, B. K., & Bearman, S. (2011). The role of efficacy and identity in science career commitment among underrepresented minority students. *Journal of Social Issues*, 67(3), 469–491. <https://doi.org/10.1111/j.1540-4560.2011.01710.x>.
- Cheong, K. H., & Koh, J. M. (2018). Integrated virtual laboratory in engineering mathematics education: Fourier theory. *IEEE Access*, 6, 58231–58243. <https://doi.org/10.1109/ACCESS.2018.2873815>.
- Chini, J. J., Madsen, A., Gire, E., Rebello, N. S., & Puntambekar, S. (2012). Exploration of factors that affect the comparative effectiveness of physical and virtual manipulatives in an undergraduate laboratory. *Physical Review Special Topics - Physics Education Research*, 8(1). <https://doi.org/10.1103/PhysRevSTPER.8.010113>.
- Cobb, S., Heaney, R., Corcoran, O., & Henderson-Begg, S. (2009). The learning gains and student perceptions of a second life virtual lab. *Bioscience Education*, 13(1), 1–9. <https://doi.org/10.3108/beej.13.5>.
- (Committee) Steering Committee on Preparing the Engineering and Technical Workforce for Adaptability and Resilience to Change, National Academy of Engineering, & National Academies of Sciences, Engineering, & Medicine. (2018). *Adaptability of the US engineering and technical workforce: proceedings of a workshop*. (K. P. Jarboe & S. Olson, Eds.). Washington, D.C.: National Academies Press. <https://doi.org/10.17226/25016>
- Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, 19(2), 272–309. <https://doi.org/10.1080/15213269.2015.1015740>.
- Dalgarno, B., Bishop, A. G., Adlong, W., & Bedgood, D. R. (2009). Effectiveness of a virtual laboratory as a preparatory resource for distance education chemistry students. *Computers and Education*, 53(3), 853–865. <https://doi.org/10.1016/j.compedu.2009.05.005>.
- Darrah, M., Humbert, R., Finstein, J., Simon, M., & Hopkins, J. (2014). Are virtual labs as effective as hands-on labs for undergraduate physics? A comparative study at two major universities. *Journal of Science Education and Technology*, 23(6), 803–814. <https://doi.org/10.1007/s10956-014-9513-9>.
- de Jong, T., Linn, M. C., & Zacharia, Z. C. (2013). Physical and virtual laboratories in science and engineering education. *Science*, 340(6130), 305–308. <https://doi.org/10.1126/science.1230579>.
- de Jong, T., Sotiriou, S., & Gillet, D. (2014). Innovations in STEM education: the go-lab federation of online labs. *Smart Learning Environments*, 1(1). <https://doi.org/10.1186/s40561-014-0003-6>.
- Driscoll, M. (2005). *Psychology of learning for instruction*. Boston: Pearson.
- Dyrberg, N. R., Treusch, A. H., & Wiegand, C. (2017). Virtual laboratories in science education: students' motivation and experiences in two tertiary biology courses. *Journal of Biological Education*, 51(4), 358–374. <https://doi.org/10.1080/00219266.2016.1257498>.
- Ekmekci, A., & Gulacar, O. (2015). A case study for comparing the effectiveness of a computer simulation and a hands-on activity on learning electric circuits. *Eurasia Journal of Mathematics, Science and Technology Education*, 11(4), 765–775. <https://doi.org/10.12973/eurasia.2015.1438a>.
- Fang, N., & Tajvidi, M. (2018). The effects of computer simulation and animation (CSA) on students' cognitive processes: A comparative case study in an undergraduate engineering course. *Journal of Computer Assisted Learning*, 34(1), 71–83. <https://doi.org/10.1111/jcal.12215>.
- Feng, Z., González, V. A., Amor, R., Lovreglio, R., & Cabrera-Guerrero, G. (2018). Immersive virtual reality serious games for evacuation training and research: a systematic literature review. *Computers and Education*, 127, 252–266. <https://doi.org/10.1016/j.compedu.2018.09.002>.
- Fogarty, L., Strimling, P., & Laland, K. N. (2011). The evolution of teaching. *Source: Evolution*, 65.
- Gardner, G. E., & Jones, M. G. (2011). Pedagogical preparation of the science graduate teaching assistant: challenges and implications. *Science Education*, 20(2), 31–41.
- Gao, Z., Liu, S., Ji, M., & Liang, L. (2011). Virtual hydraulic experiments in courseware: 2D virtual circuits and 3D virtual equipments. *Computer Applications in Engineering Education*, 19(2), 315–326. <https://doi.org/10.1002/cae.20313>.
- Gobaw, G. F. (2016). Analysis of undergraduate biology laboratory manuals. *International Journal of Biology Education*, 5(1). <https://doi.org/10.20876/ijobed.99404>.
- Goudsouzian, L. K., Riola, P., Ruggles, K., Gupta, P., & Mondoux, M. A. (2018). Integrating cell and molecular biology concepts: comparing learning gains and self-efficacy in corresponding live and virtual undergraduate laboratory experiences. *Biochemistry and Molecular Biology Education*, 46(4), 361–372. <https://doi.org/10.1002/bmb.21133>.
- Gunn, T., Jones, L., Bridge, P., Rowntree, P., & Nissen, L. (2018). The use of virtual reality simulation to improve technical skill in the undergraduate medical imaging student. *Interactive Learning Environments*, 26(5), 613–620. <https://doi.org/10.1080/10494820.2017.1374981>.
- Hawkins, I., & Phelps, A. J. (2013). Virtual laboratory vs. traditional laboratory: which is more effective for teaching electrochemistry?

- Chemistry Education Research and Practice*, 14(4), 516–523. <https://doi.org/10.1039/c3rp00070b>.
- Healey, M., & Jenkins, A. (2009). Developing undergraduate research and inquiry. In *AdvanceHE* Retrieved from <https://www.advance-he.ac.uk/knowledge-hub/developing-undergraduate-research-and-inquiry>.
- Hofstein, A. (2004). The laboratory in chemistry education: thirty years of experience with developments, implementation, and research. *Chemical Education Research and Practice*, 5(3), 247–264. <https://doi.org/10.1039/B4RP90027H>.
- Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: foundations for the twenty-first century. *Science Education*, 88(1), 28–54. <https://doi.org/10.1002/sce.10106>.
- Holden, J.P., Lander, E. (2012) *Report to the president: engage to excel: producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Executive Office of the President President’s Council of Advisors on Science and technology.
- Jensen, L., & Konradsen, F. (2017). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies*, 23(4), 1–15. <https://doi.org/10.1007/s10639-017-9676-0>.
- Jones, N. (2018). Simulated labs are booming. *Nature*, 562(7725), S5–S7. <https://doi.org/10.1038/d41586-018-06831-1>.
- Kang, J., Lindgren, R., & Planey, J. (2018). Exploring emergent features of student interaction within an embodied science learning simulation. *Multimodal Technologies and Interaction*, 2(3). <https://doi.org/10.3390/mti2030039>.
- Kempton, C. E., Weber, K. S., & Johnson, S. M. (2017). Method to increase undergraduate laboratory student confidence in performing independent research. *Journal of Microbiology & Biology Education*, 18(1), 1–4. <https://doi.org/10.1128/jmbe.v18i1.1230>.
- Kester, L., Kirschner, P. A., van Merriënboer, J. J. G., & Baumer, A. (2001). Just-in-time information presentation and the acquisition of complex cognitive skills. *Computers in Human Behavior*, 17(4), 373–391. [https://doi.org/10.1016/S0747-5632\(01\)00011-5](https://doi.org/10.1016/S0747-5632(01)00011-5).
- Khan, K. S., Kunz, R., Kleijnen, J., & Antes, G. (2003). Five steps to conducting a systematic review. *Journal of the Royal Society of Medicine*, 96. Retrieved from [http://www.ncbi.nlm.nih.gov/entrez/query/\(3\)](http://www.ncbi.nlm.nih.gov/entrez/query/(3)), 118–121.
- Koh, C., Tan, H., Kim Tan, C., Fang, L., Fong, F., Kan, D., Lye, S., & Wee, M. (2010). Investigating the effect of 3D simulation-based learning on the motivation and performance of engineering students. *Journal of Engineering Education*, 99(3), 237–251. <https://doi.org/10.1002/j.2168-9830.2010.tb01059.x>.
- Krippendorff, K. (2012). *Content analysis: an introduction to its methodology*. Los Angeles: SAGE Publications, Inc..
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Los Angeles: SAGE Publications, Inc..
- Ma, J., & Nickerson, J. V. (2006). Hands-on, simulated, and remote laboratories: a comparative literature review. *ACM Computing Surveys*, 38(3), 1. <https://doi.org/10.1145/1132960.1132961>.
- Makransky, G., & Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. *Educational Technology Research and Development*, 66(5), 1141–1164. <https://doi.org/10.1007/s11423-018-9581-2>.
- Makransky, G., Mayer, R., Nøremølle, A., Cordoba, A. L., Wandall, J., & Bonde, M. (2020). Investigating the feasibility of using assessment and explanatory feedback in desktop virtual reality simulations. *Educational Technology Research and Development*, 68(1), 293–317. <https://doi.org/10.1007/s11423-019-09690-3>.
- Makransky, G., Mayer, R. E., Veitch, N., Hood, M., Christensen, K. B., & Gadegaard, H. (2019). Equivalence of using a desktop virtual reality science simulation at home and in class. *PLoS ONE*, 14(4), 1–14. <https://doi.org/10.1371/journal.pone.0214944>.
- Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2017). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, 60, 225–236. <https://doi.org/10.1016/j.learninstruc.2017.12.007>.
- Makransky, G., Thisgaard, M. W., & Gadegaard, H. (2016). Virtual simulations as preparation for lab exercises: assessing learning of key laboratory skills in microbiology and improvement of essential non-cognitive skills. *PLoS One*, 11(6), e0155895. <https://doi.org/10.1371/journal.pone.0155895>.
- Martín-Gutiérrez, J., Mora, C. E., Añorbe-Díaz, B., & González-Marrero, A. (2017). Virtual technologies trends in education. *Eurasia Journal of Mathematics, Science and Technology Education*, 13(2), 469–486. <https://doi.org/10.12973/eurasia.2017.00626a>.
- Mavromanolakis, G., Lazoudis, A., & Sotiriou, S. A. (2014). Diffusion of inquiry-based science teaching methods and practices across Europe: experience and outcomes from the “pathway”, a project supported by the 7th Framework Programme of the European Commission. In *2014 IEEE 14th International Conference on Advanced Learning Technologies* (pp. 734–736). IEEE. <https://doi.org/10.1109/ICALT.2014.214>.
- Mayer, R. E. (2014). *The Cambridge handbook of multimedia learning, second edition. The Cambridge Handbook of Multimedia Learning*. 2nd Edition: Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369>.
- Michel, N., Cater, J. J., & Varela, O. (2009). Active versus passive teaching styles: an empirical study of student learning outcomes. *Human Resource Development Quarterly*, 20(4), 397–418. <https://doi.org/10.1002/hrdq.20025>.
- Mikropoulos, T. A., & Natsis, A. (2011). Educational virtual environments: a ten-year review of empirical research (1999–2009). *Computers and Education*, 56(3), 769–780. <https://doi.org/10.1016/j.compedu.2010.10.020>.
- Nolen, S. B., & Koretsky, M. D. (2018). Affordances of virtual and physical laboratory projects for instructional design: impacts on student engagement. *IEEE Transactions on Education*, 61(3), 226–233. <https://doi.org/10.1109/TE.2018.2791445>.
- (NRC) National Research Council. (2006). *America’s lab report: investigations in high school science*. National Academies Press.
- (NRC) National Research Council. (2012). *A framework for K-12 science education: practices, crosscutting concepts, and core ideas*. Washington, DC, USA: National Academy Press.
- Ogbuanya, T. C., & Onele, N. O. (2018). Investigating the effectiveness of desktop virtual reality for teaching and learning of electrical/electronics technology in universities. *Computers in the Schools*, 35(3), 226–248. <https://doi.org/10.1080/07380569.2018.1492283>.
- Olympiou, G., & Zacharia, Z. C. (2012). Blending physical and virtual manipulatives: an effort to improve students’ conceptual understanding through science laboratory experimentation. *Science Education*, 96(1), 21–47. <https://doi.org/10.1002/sce.20463>.
- Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: a practical guide*. Malden, MA: Blackwell.
- Polly, P., Marcus, N., Maguire, D., Belinson, Z., & Velan, G. M. (2014). Evaluation of an adaptive virtual laboratory environment using Western blotting for diagnosis of disease. *BMC Medical Education*, 14(1). <https://doi.org/10.1186/1472-6920-14-222>.
- Potkonjak, V., Gardner, M., Callaghan, V., Mattila, P., Guetl, C., Petrović, V. M., & Jovanović, K. (2016). Virtual laboratories for education in science, technology, and engineering: a review. *Computers and Education*, 95, 309–327. <https://doi.org/10.1016/j.compedu.2016.02.002>.
- Reid, N., & Shah, I. (2007). The role of laboratory work in university chemistry. *Chemical Education Research and Practice*, 8(2), 172–185. <https://doi.org/10.1039/B5RP90026C>.
- Reyes-Aviles, F., & Aviles-Cruz, C. (2018). Handheld augmented reality system for resistive electric circuits understanding for undergraduate

- students. *Computer Applications in Engineering Education*, 26(3), 602–616. <https://doi.org/10.1002/cae.21912>.
- Riva, G., Davide, F., & Ijsselstein, W. A. (2003). Being there: The experience of presence in mediated environments. In *Being There: Concepts, Effects and Measurements of User Presence in Synthetic Environments* (Vol. 5). Amsterdam: IOS Press.
- Rush, S., Acton, L., Tolley, K., Marks-Maran, D., & Burke, L. (2010). Using simulation in a vocational programme: does the method support the theory? *Journal of Vocational Education and Training*, 62(4), 467–479. <https://doi.org/10.1080/13636820.2010.523478>.
- Ryker, K. D., & McConnell, D. A. (2017). Assessing inquiry in physical geology laboratory manuals. *Journal of Geoscience Education*, 65(1), 35–47. <https://doi.org/10.5408/14-036.1>.
- Schott, C., & Marshall, S. (2018). Virtual reality and situated experiential education: a conceptualization and exploratory trial. *Journal of Computer Assisted Learning*, 34(6), 843–852. <https://doi.org/10.1111/jcal.12293>.
- Secomb, J., Mckenna, L., & Smith, C. (2012). The effectiveness of simulation activities on the cognitive abilities of undergraduate third-year nursing students: a randomised control trial. *Journal of Clinical Nursing*, 21(23–24), 3475–3484. <https://doi.org/10.1111/j.1365-2702.2012.04257.x>.
- Singer, S. R., Hilton, M. L., & Schweingruber, H. A. (2005). America's lab report: Investigations in High School Science. *America's Lab Report: Investigations in High School Science*. <https://doi.org/10.17226/11311>.
- Stains, M., & Vickrey, T. (2017). Fidelity of implementation: an overlooked yet critical construct to establish effectiveness of evidence-based instructional practices. *CBE Life Sciences Education*, 16(1), m1. <https://doi.org/10.1187/cbe.16-03-0113>.
- Tatli, Z., & Ayas, A. (2013). Effect of a virtual chemistry laboratory on students' achievement. *Educational Technology & Society*, 16(1), 159–170 Retrieved from http://www.ifets.info/journals/16_1/14.pdf.
- Tetour, Y., Boehringer, D., & Richter, T. (2011). Integration of virtual and remote experiments into undergraduate engineering courses.
- Toth, E. E. (2016). Analyzing “real-world” anomalous data after experimentation with a virtual laboratory. *Educational Technology Research and Development*, 64(1), 157–173. <https://doi.org/10.1007/s11423-015-9408-3>.
- Ural, E. (2016). The effect of guided-inquiry laboratory experiments on science education students' chemistry laboratory attitudes, anxiety and achievement. *Journal of Education and Training Studies*, 4(4), 217–227. <https://doi.org/10.11114/jets.v4i4.1395>.
- Uribe, M. D. R., Magana, A. J., Bahk, J. H., & Shakouri, A. (2016). Computational simulations as virtual laboratories for online engineering education: a case study in the field of thermoelectricity. *Computer Applications in Engineering Education*, 24(3), 428–442.
- Vrellis, I., Avouris, N., & Mikropoulos, T. A. (2016). Learning outcome, presence and satisfaction from a science activity in second life. *Australasian Journal of Educational Technology*.
- Wan, T., Geraets, A. A., Doty, C. M., Saitta, E. K. H., & Chini, J. J. (2020). Characterizing science graduate teaching assistants' instructional practices in reformed laboratories and tutorials. *International Journal of STEM Education*, 7(1), 30. <https://doi.org/10.1186/s40594-020-00229-0>.
- Wells, J. D., Campbell, D. E., Valacich, J. S., & Featherman, M. (2010). The effect of perceived novelty on the adoption of information technology innovations: a risk/reward perspective. *Decision Sciences*, 41(4), 813–843. <https://doi.org/10.1111/j.1540-5915.2010.00292.x>.
- Williams, D. (2010). The mapping principle, and a research framework for virtual worlds. *Communication Theory*, 20(4), 451–470. <https://doi.org/10.1111/j.1468-2885.2010.01371.x>.
- Zhu, H., Yang, Z., Xiong, Y., Wang, Y., & Kang, L. (2018). Virtual emulation laboratories for teaching offshore oil and gas engineering. *Computer Applications in Engineering Education*, 26(5), 1603–1613. <https://doi.org/10.1002/cae.21977>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.