Are virtual physiology laboratories effective for student learning? A systematic review

Xinyu Zhang, Dawoud Al-Mekhled, and Julia Choate
Department of Physiology, Monash University, Victoria, Australia

Abstract
It is unclear if the transition from traditional, in-person physiology laboratories to virtual alternatives has educational impacts on students. This study used a systematic review to critically evaluate research papers that investigated the effectiveness of virtual physiology laboratories for student learning. Eleven studies, retrieved from the Education Resources Information Center (ERIC) and Ovid MEDLINE databases, were selected for inclusion in this review, based on predetermined eligibility criteria. Subsequently, the studies went through a power analysis for potential biases before their results were synthesized and analyzed. This systematic review found that virtual physiology laboratories are effective for students’ learning of concepts. However, it was inconclusive as to whether virtual physiology laboratories are effective for students’ motivation for learning and learning of technical skills. It was found that blended models of virtual laboratories are at least as effective as in-person laboratories for conceptual learning. Overall, this systematic review provides useful insights for educators regarding the educational impacts of implementing virtual laboratories into the physiology curriculum and suggests research models for future evaluation of virtual laboratories.

physiology; simulation; undergraduate; virtual laboratory

INTRODUCTION
The Problem: A Forced Shift to Virtual Laboratories
In 2020, the coronavirus disease 2019 (COVID-19), caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), led to a pandemic. In response to this pandemic, the Australian Government declared a human biosecurity emergency and imposed lockdown measures to contain the transmission of the virus (1). This included social distancing, one of the most effective strategies to limit the spread of the virus through human-to-human transmission (2). To comply with these social distancing measures, the delivery of in-person teaching models at universities, including laboratories, was forced to shift to remote, or virtual, alternatives. Laboratories, a critical component of undergraduate science education (3, 4), have thus been severely impacted by the COVID-induced restrictions (5, 6). For the physiology discipline, traditional, in-person laboratories play a central role to support students’ understanding of physiology concepts and the development of research and technical skills (7–9). In particular, the social interactions and the teamwork during in-person laboratories support collaborative learning (3). As a result of university lockdowns, the remote delivery of virtual physiology laboratories is thus likely to have educational impacts on student learning.

Historical Development of Virtual Laboratories
A virtual laboratory is any online environment that is based on interactive learning either individually or in groups, allowing students to explore topics in an asynchronous manner that has no immediate physical reality (10). Virtual laboratories have been developed since the 1990s, where educators used readily accessible technologies such as the compact disk read-only memory (CD-ROM) to design virtual laboratories that could be accessed by students (11). Contemporary designs of virtual laboratories utilize network-based configurations to create virtual simulations for manipulation. In these virtual laboratories, participants manipulate virtual apparatus to perform experiments in computer-generated simulations. For example, virtual reality computer programs have been used to teach robotics to engineering students (12). Additionally, web packages, which are website-based platforms with interactive components, have been used to teach undergraduate students about the cardiovascular system (13). Furthermore, virtual simulations of animal experiments have been used for teaching physiology (14–16). In the last decade, there has been a gradual shift of conventional physical laboratories toward virtual alternatives, to either fully replace or support the physical laboratories (17). There is also an increasing number of university disciplines teaching via web-based instruments and more university courses that are being delivered entirely online (18).
Pre-COVID Reasons for the Use of Virtual Laboratories

Before the COVID-19 pandemic, there were four main reasons that promoted the transition of physical, in-person, laboratories to virtual alternatives. First, conducting physical laboratories is financially demanding. This can be attributed to the provision of advanced instruments, space, the recruitment of professional personnel, the maintenance of equipment, and the rapidly increasing student population (2, 19). Second, animal-based laboratories can be associated with ethical concerns, therefore, limiting the scope of experiments that educators can employ for students (20). In comparison, using virtual animal model simulations reduces the ethical dilemmas and broadens the types of experiments that can be conducted. Moreover, as compared with in-person laboratories, virtual laboratories exhibit higher levels of efficiency (21) and safety (22, 23). Lastly, prior studies have reported that virtual learning strategies such as simulations and instructional videos can be used to support the physical laboratories in enhancing students’ learning (7, 24–26). For these reasons, the transition from traditional in-person laboratories to virtual alternatives is likely to persist in the higher education sector. However, in-person laboratories, which provide tactile hands-on experiences, are undeniably different in nature to the virtual laboratories, which have no physical reality. Therefore, this transition is likely to impact student learning. While the in-person laboratories are unquestionably effective for student learning (27), it is unclear as to whether virtual laboratories support student learning.

Student Learning for Physiology Laboratories

Due to the increasing use of virtual laboratories, particularly in response to the COVID-19 pandemic, it is critical to determine if they are effective for student learning. For this study, effective student learning was defined as student achievement of the (laboratory) learning outcomes (28). Laboratory learning outcomes cover both physiology content/concepts and skills or competencies. The type of student learning from laboratories was divided into three categories (29):

1) Conceptual learning: the laboratory reinforces or enhances the understanding of physiology concepts taught in other teaching modes, such as lectures and workshops;
2) Motivational learning: the laboratory motivates engagement with scientific processes, scientific exploration and active learning from failures (in experimental design, for example); and
3) Technical learning or research skills development: the laboratory allows the practice and learning of technical skills such as using scientific equipment and other research skills such as conducting data analysis, interpretation, and presentation.

Systematic Review for Evaluating the Research Literature

In the physiology discipline, despite a number of primary research articles about virtual laboratories (for example, see Refs. 14, 24, 30), there has been no critical evaluation of these papers. This study used a systematic review to critically appraise the existing literature, to determine if virtual physiology laboratories are effective for student learning. A systematic review is a robust process that critically analyses previous literature through explicit methodology, minimizing biases and enabling the synthesis and generation of comprehensive conclusions (31–33). The systematic review was expected to identify articles that highlight the educational impacts of virtual physiology laboratories on student learning and to reveal gaps in the research literature, guiding the direction of future research.

METHODS

Study Design

This systematic review was conducted following the Preferred Reporting Items for Systematic Reviews & Meta-Analyses (PRISMA) guidelines (34), a framework that allows the design of well-defined methodology to evaluate the existing literature from an unbiased and objective viewpoint (31).

Search Strategy

The electronic database search included the Education Resources Information Center (ERIC) and Ovid MEDLINE. ERIC was selected because of its extensive collection of educational research that is expected to include a large number of research publications for this systematic review. Ovid MEDLINE was selected because it is an authoritative database that includes a comprehensive collection of medical-related research that is relevant to physiology. The scope of the literature search was limited by five categories, including 1) virtual; 2) laboratory; 3) undergraduate students; 4) physiology discipline(s); and 5) educational benefits. Keywords and subject headings from each category were used to conduct the literature search (see Table 1).

The search was conducted on the 21st of September 2020, using the following steps: 1) for each category, the keywords were searched individually; 2) for each category, the subject headings were searched individually; 3) for each category, the search results of all keywords and subject headings were combined with the operator “OR”; 4) the search results from all categories were combined with the operator “AND”; and 5) the search was limited to papers from 2005 as this aligned with the period for significant development of virtual laboratory technologies.

Eligibility Criteria

Database search results were imported into an EndNote library, including the portable document format (pdf) of the full-text articles. Duplicates were then removed, and initial screening was completed by title and abstract. This was followed by screening of the full-text articles. Exclusion of papers was completed by the three researchers independently, followed by consensus for exclusion by the researchers. The relevant inclusion criteria are identified below, and reasons for exclusions are noted in the PRISMA flow chart (Fig. 1):

1) Language: the literature was presented using the English language;
2) Full text;
3) Peer-reviewed articles;
4) Physiology discipline: the function of organisms, including subdisciplines such as cell physiology, neurophysiology, endocrinology, immunology, exercise physiology, pathophysiology, comparative physiology, cardiovascular physiology, respiratory physiology, psychophysiology, renal physiology, and systems physiology;
5) Undergraduate: tertiary/undergraduate education, after secondary education and before postgraduate education; and
6) Critical evaluation of virtual physiology laboratories: the literature conducted formal evaluation(s) of the virtual instrument used in the study from any of the following aspects: conceptual learning, motivational learning, and technical learning.

A Google Excel spreadsheet was created to systematically screen the literature. The spreadsheet included the basic information (title, citation), the key findings, and the eligibility criteria to select the relevant papers. The screening process began by checking the titles and the abstracts of the articles against the eligibility criteria. Following the elimination of irrelevant articles by titles and abstracts, the full texts of the remaining papers were assessed against the eligibility criteria. After the screening process, a team consensus was reached for the eligibility criteria, before finalizing the relevant papers for analysis.

### Assessment of the Risk of Bias

As the included studies comprised a mixture of both quantitative and qualitative studies, the conventional data analysis tools were not applicable to this review. Therefore, the risk of bias in individual studies was evaluated by using a novel additive scoring system. Firstly, nine bias/error-assessing criteria were nominated to evaluate the risk of bias/error of the studies. This included 1) randomization; 2) sample size; 3) voluntary participation; 4) presence of control group(s); 5) group sizes; 6) consistency across groups; 7) response rate; 8) type of evaluation; and 9) data significance (see Table 2). A score was allocated to each criterion, and then the scores were added for all of the criteria to produce a final score (out of 5) that indicates the power of the study. Criteria 1, 4, 5, and 6 were considered together as a lack of control group(s) simultaneously invalidates the other three criteria. Criteria 3 and 7 were considered together as although volunteerism can increase the risk of volunteer bias, a high response rate suggests a high level of inclusion of the sample with a lower risk of sampling bias; therefore, each criterion was allocated 0.5 points. Evaluation of the literature involved examining the fulfillment of the criteria, and the respective points were scored. The threshold of valid sample sizes used in this study was 50 (35), and the threshold of valid response rate used in this study was 60% (36). Table 2 summarizes the criteria, their definitions, and their score allocations.

### Table 1. Keywords and subject headings for the literature search

| Categories            | Keywords†                                                                 | Subject Headings (Ovid MEDLINE)                                                                 | Subject Headings (ERIC)                                                                 |
|-----------------------|---------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------|
| Virtual               | Virtual, Simul*, Blended learning, Blended teaching, Computer*, Online    | Computer simulation (methods, standards), Virtual reality, Simulation training (methods, standards), Online systems (education, instrumentation methods, standards) | Virtual classrooms, Virtual universities, Computer uses in education, Computer simulation, Online systems, Simulation |
| Laboratory            | Lab* based, Lab* class*, Lab* setting*, Practical class*, Practical lab*  | Laboratories (instrumentation, methods, standards)                                             | Laboratory techniques, Laboratory training, Laboratory equipment                        |
| Undergraduate students| Undergraduate*, Universities (education, standards)                        |                                                                                               | College students, Undergraduate students                                                 |
| Physiology discipline†| Physiology, Animal*, Neurophysiology, Endocrinology, Immunology, Pathophysiology, Psychophysiology | Physiology (analysis, classification, drug effects, education, ethics, instrumentation, methods, standards, organization and administration, statistics and numerical data, supply and distribution, surgery, trends) | Exercise physiology, Exercise (physiology), Hearing (physiology), Psychophysiology, Animals, Physiology |
| Educational benefits  | Benefit*, Advantage*, Effective*, Education*, Pitfall*, Limitation*, Disadvantage*, Evaluat* | Evaluation study, Program evaluation (methods, standards, statistics and numerical data, trends) | Educational opportunities, Educational assessment, Educational benefits, Program evaluation |

†Same keywords were used in both Ovid MEDLINE and Education Resources Information Center (ERIC) during the literature search, whereas the subject headings were restricted to the unique thesaurus of each database, therefore, suitable subject headings were selected from each database. The asterisks (*) allow the retrieval of all words that have the same starting letters and variable ending letters. ‡Including the keyword "physiology" will retrieve all terms including these letters; therefore, subdisciplines such as "respiratory physiology" were not included as keywords.
Data Items

The mode of delivery is defined as the educational configuration of the laboratory activity. For example, a blended mode of delivery involves both the components of the physical laboratory as well as the components of the virtual laboratory. In contrast, a standalone virtual mode of delivery includes no physical laboratory. A critical evaluation of virtual tools refers to the formal assessment of the virtual instrument used in the study for student learning from any of the following aspects: conceptual learning, motivational learning, and technical learning.

Ethics Statement

Institutional Review Board approval was not necessary for this protocol as it was a systematic review, using data from publicly accessible data.

RESULTS

Literature Search

The literature search retrieved 57 papers from ERIC and 58 papers from Ovid MEDLINE, summing up to 115 papers in total (Fig. 1). After removal of 11 duplicates, the abstracts for the remaining 104 papers were screened against the eligibility criteria, following the order of English, full-text, primary journals, physiology discipline(s), undergraduate level, virtual laboratory, critical evaluation. All papers were in English and were primary studies. This resulted in the exclusion of 75 papers. When the full text for the remaining 29 papers was assessed, 16 papers were excluded (see Fig. 1 for the breakdown of the exclusion criteria). A total of 13 papers were included for subsequent evaluation. Detailed information about the number of papers excluded and the corresponding reasons for exclusion at each stage of screening are shown in Fig. 1.

Risk of bias assessment.

Based on the novel additive scoring system used, 2 out of the 13 included studies (papers) had low power and high risk of bias/error. A revisit of these studies revealed deficiencies in their study designs; therefore, both studies were removed from this review.

Data Extraction from the Included Studies

The participants, study design, research aims, laboratory, and key findings for the final 11 included studies in this systematic review are presented in Table 3.
Demographics.
All of the final included studies for this systematic review were conducted at different universities. A large proportion of the studies were from the United States (45.5%), followed by China (18.2%), Australia (9.1%), Brazil (9.1%), Spain (9.1%), and Turkey (9.1%). The participants of the studies were largely undergraduate students including medical (45.5%), exercise physiology (18.2%), and nursing students (9.1%). The remaining participants were from science degrees (27.3%), including biomedical and biological science.

Study designs.
Three of the eleven studies conducted randomized controlled trials, three adopted a quasieperimental study design, two adopted a within-subject design, two adopted a crossover study design, and one adopted a cross-sectional study design (see Table 3).

Virtual tools and modes of delivery.
The virtual tools and modes of delivery for the 11 papers are summarized in Table 3. For the virtual laboratories used, 36.4% were virtual simulations (usually an imitation of the physical laboratory), 27.3% were web-based laboratory courses (website with interactive activities), 27.3% were video demonstrations (video displaying experimental procedures and outcomes), and 9.1% were computer software (inquiry-based system that allows students to perform experiments). Of the 11 studies, 5 were delivered in an asynchronous manner (7, 24, 37, 38, 42), 5 were delivered in a synchronous manner (13, 16, 39–41), and 1 was delivered via both modes (25). To test these virtual laboratories, 63.6% of the studies conducted the

---

Table 2. Method for the assessment of potential bias/error within studies

| No | Criteria | Definition | Evaluation (Score Allocation) |
|----|----------|------------|-------------------------------|
| 1  | Randomization | The random allocation of participants to the experimental/control groups. | Yes (0.25): The presence of randomization indicates a lower risk of selection bias. No (0): The absence of randomization indicates a higher risk of selection bias. N/A (0): Not applicable for studies with only 1 group. |
| 2  | Sample size | A minimum of 50 participants are required to establish a valid evaluation. | Large (1): Large sample size indicates a lower risk of sampling bias. Small (0): Small sample size indicates a higher risk of sampling bias. |
| 3  | Voluntary participation | The presence of volunteerism in the study, which includes volunteering to participate in the study or provide data for the study. | Yes (0): Using volunteers for a study indicates a higher risk of volunteer bias. No (0.5): Using nonvolunteers for a study indicates a lower risk of volunteer bias. |
| 4  | Presence of control group(s) | A group without the implementation of the target intervention. It is compared with the experimental group, where the target intervention is implemented. | Yes (0.25): The presence of control group(s) indicates a lower risk of placebo effect. No (0): The absence of control group(s) indicates a higher risk of placebo effect. |
| 5  | Group sizes | The number of participants per group is balanced. | Yes (0.25): Balanced group sizes indicate a lower risk of selection bias, sampling bias, and type I errors. No (0): Unbalanced group sizes indicate a higher risk of selection bias, sampling bias, and type I errors. N/A (0): Not applicable for studies with only 1 group. |
| 6  | Consistency across groups | Demographics of participants across the groups are consistent without significant variation between them. | Yes (0.25): Consistent demographics across groups indicates a lower risk of selection bias and sampling bias. No (0): Inconsistent demographics across groups indicates a higher risk of selection bias and sampling bias. Not evident (0): The risks of biases cannot be determined. N/A (0): Not applicable for studies with only 1 group. |
| 7  | Response rate | The proportion of participant response compared with the total number of participants in the cohort/subject. | High (0.5): High response rate indicates a lower risk of nonresponse bias. Low (0): Low response rate indicates a higher risk of nonresponse bias. |
| 8  | Type of evaluation | Survey/Questionnaire vs. Test score. | Survey/Questionnaire (0.5): Evaluation using surveys/questionnaires can be subjected to response biases. Test score (1): Evaluation using test score eliminates response biases. |
| 9  | Data significance | The significance of data collected is investigated. | Yes (1): Investigating the data significance indicates a lower risk of interpretive bias. No (0): Not investigating the data significance indicates a higher risk of interpretive bias. |
Table 3. Study characteristics

| Reference                        | Participants                      | Study Design                        | Research Aims                                                                 | Laboratory                                                                 | Key Findings                                                                                                                                   |
|----------------------------------|-----------------------------------|-------------------------------------|-------------------------------------------------------------------------------|-------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------|
| Moreno-Ger et al. (37)           | 143 medical undergraduates        | Randomized control trial design     | To investigate:                                                               | Aims:                                                                           | The EG perceived that the procedures were easier to understand and displayed a higher level of precision in determining the hematocrit. The EG reported a more positive experience. |
|                                  |                                   | Experimental group (EG), n = 66, used the virtual laboratory simulation prior to the physical laboratory | If having prelab virtual simulation decreases the perceived difficulty for the physical laboratory. | Mode of delivery: Virtual Synchronous | EG made stronger improvements between the pretest and posttest than CG. However, due to significant inconsistency between CG and EG, the comparison between CG and EG was considered invalid. Nonetheless, the improvement within the EG was significant. |
|                                  |                                   | Control group (CG), n = 77, attended the physical laboratory only | If having prelab virtual simulation increases the experimental precision during physical laboratory. | Virtual tool: Virtual laboratory simulation | The high-fidelity simulator was more effective compared with the low-fidelity simulator in supporting students’ conceptual understanding in heart and lung sounds. |
|                                  |                                   | Student perception: 5-point Likert survey | The students’ attitudes toward a prelab virtual simulation. | Mode of delivery: Blended Asyncronous | The high-fidelity simulator was more effective compared with the low-fidelity simulator in supporting students’ conceptual understanding in heart and lung sounds. |
| Mutlu et al. (39)                | 71 undergraduates from a nursing faculty | Quasi-experimental study design    | To determine the effects of high- and low-fidelity simulators on student nurses’ learning of heart and lung sounds. | Aim:                                                                           | The high-fidelity simulator was more effective compared with the low-fidelity simulator in supporting students’ conceptual understanding in heart and lung sounds. |
|                                  |                                   | EG, n = 104, CG, n = 69, Pretest and posttest results were compared. | Aim: To develop students’ skills to accurately identify heart and lung sounds. | Virtual tool: Virtual laboratory simulation | The high-fidelity simulator was more effective compared with the low-fidelity simulator in supporting students’ conceptual understanding in heart and lung sounds. |
| Wang et al. (16)                 | 63 undergraduates (Bachelor of Medicine and Bachelor of Surgery) | Randomized control trial design     | To investigate the impact of the virtual laboratory simulation in supporting students’ learning and perceptions of neurophysiology. | Mode of delivery: Virtual Synchronous | Both the live tissue laboratory and the blended laboratory were more effective in supporting conceptual learning than the virtual laboratory. The live tissue, blended, and virtual laboratories were equally effective in supporting students’ technical learning. Delivering animal experiments via virtual means was deemed less attractive compared with the physical and the blended delivery modes. The blended laboratory course was effective in enhancing |
|                                  |                                   | High-fidelity group (HFS), n = 36, used Nasco Smartscope Simulator | To assess the validity of the virtual laboratory simulation as a useful pedagogical tool for neurophysiology. | Virtual tool: Virtual laboratory simulation | Both the live tissue laboratory and the blended laboratory were more effective in supporting conceptual learning than the virtual laboratory. The live tissue, blended, and virtual laboratories were equally effective in supporting students’ technical learning. Delivering animal experiments via virtual means was deemed less attractive compared with the physical and the blended delivery modes. The blended laboratory course was effective in enhancing |
|                                  |                                   | Low-fidelity group (LFS), n = 35, used computer-based simulator | To provide basic training of technical learning. | Mode of delivery: Virtual Synchronous | Both the live tissue laboratory and the blended laboratory were more effective in supporting conceptual learning than the virtual laboratory. The live tissue, blended, and virtual laboratories were equally effective in supporting students’ technical learning. Delivering animal experiments via virtual means was deemed less attractive compared with the physical and the blended delivery modes. The blended laboratory course was effective in enhancing |
|                                  |                                   | Pretest and posttest results were compared. | Aim: To develop students’ skills to accurately identify heart and lung sounds. | Virtual tool: Low-fidelity simulator (computer-based simulator) | Both the live tissue laboratory and the blended laboratory were more effective in supporting conceptual learning than the virtual laboratory. The live tissue, blended, and virtual laboratories were equally effective in supporting students’ technical learning. Delivering animal experiments via virtual means was deemed less attractive compared with the physical and the blended delivery modes. The blended laboratory course was effective in enhancing |
| Chen et al. (24)                | 173 medical undergraduates        | Within-subject design               | To investigate if the blended laboratory | Mode of delivery: Virtual and blended Synchronous | Both the live tissue laboratory and the blended laboratory were more effective in supporting conceptual learning than the virtual laboratory. The live tissue, blended, and virtual laboratories were equally effective in supporting students’ technical learning. Delivering animal experiments via virtual means was deemed less attractive compared with the physical and the blended delivery modes. The blended laboratory course was effective in enhancing |
| Reference         | Participants                                                                 | Study Design                                                                 | Research Aims                                                                 | Laboratory                                                                 | Key Findings                                                                 |
|-------------------|-------------------------------------------------------------------------------|------------------------------------------------------------------------------|-------------------------------------------------------------------------------|----------------------------------------------------------------------------|--------------------------------------------------------------------------------|
| Dantas and Kemm   | 34 teachers who taught the blended laboratory course                           | traditional laboratory course in 2015 and a blended laboratory course in 2016. | Average scores for laboratory quizzes were compared. Student and teacher perception surveys were conducted. | course supports students' understanding in threshold concepts and improves their abilities of self-learning, understanding, and problem solving. | evaluated multiple laboratories. There is no explicit description of the laboratories. Virtual tool: Blended laboratory course, including micro videos. Mode of delivery: Blended. Asynchronous. Students preferred the blended laboratory course compared with the traditional methods. Teachers had ascertained the effectiveness of the blended model of virtual physiology laboratories in supporting students' conceptual and motivational learning. The majority of the students did not agree that the virtual physiology laboratory motivated them to learn. The e-learning marks correlated with final examination marks. |
| Dobson (40)       | 31 undergraduates studying a physiology course in a Science degree            | Within-subject design Students completed e-Learning activities as prelab exercises. | To investigate student use of e-learning in the laboratory-based course. To evaluate the learning outcomes that are enhanced by active learning facilitated by the e-learning component of the course. | To investigate student use of e-learning in the laboratory-based course. To evaluate the learning outcomes that are enhanced by active learning facilitated by the e-learning component of the course. | Aim: Not applicable as the study evaluated multiple laboratories. There is no explicit description of the laboratories. Virtual tool: e-Learning platform, with active learning promoted through hypothesis testing and predictions for physical laboratories. Mode of delivery: Blended. Asynchronous. The majority of the students did not agree that the virtual physiology laboratory motivated them to learn. The e-learning marks correlated with final examination marks. |
| Durand et al. (41)| 25 undergraduates from the Department of Applied Physiology and Kinesiology | Crossover study design. EG 1, n = 12, completed the traditional exercise physiology laboratory before undertaking a knowledge assessment. Next, the group completed the Virtual Physiology of Exercise Laboratory (VPEL) program before undertaking a knowledge assessment. EG 2, n = 13, completed the VPEL program before undertaking a knowledge assessment. Next, the group completed the traditional exercise physiology laboratory before undertaking a knowledge assessment. | To compare student learning (experimental setup, conduct, and data analysis) from the VPEL program with that from traditional exercise physiology laboratory activities. To determine if the order in which the different types of laboratories were delivered affects students' conceptual understanding in exercise physiology. | To compare student learning (experimental setup, conduct, and data analysis) from the VPEL program with that from traditional exercise physiology laboratory activities. To determine if the order in which the different types of laboratories were delivered affects students' conceptual understanding in exercise physiology. | Aim: To enhance students' conceptual understanding in lactate and ventilatory threshold and maximal oxygen consumption during exercise. No significant differences between the 2 groups for experimental assessment. Virtual tool: VPEL simulation. Mode of delivery: Virtual. Synchronous. Both the animal-using laboratories and the virtual physiology laboratory were well received by the students. The e-learning marks correlated with final examination marks. |
| Reference         | Participants                                      | Study Design                      | Research Aims                                                                                      | Laboratory                                                                 | Key Findings                                                                                          |
|-------------------|---------------------------------------------------|-----------------------------------|----------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------|
| Elmer et al. (25) | 33 undergraduates from the Department of Kinesiology and Integrative Physiology | Crossover study design            | To compare student performance and perceptions between blended and traditional laboratories in an undergraduate exercise physiology. |                                                                                       | The blended model of virtual physiology laboratory was equally effective as the traditional living tissue experiments in improving students' performances. Students viewed the use of animals as more effective to stimulate their motivational learning. Students agreed that both the animal-using laboratories and the virtual physiology laboratories were effective to support technical learning, however, they differed in terms of the skill sets that they developed. |
| Gopal et al. (13) | 165 undergraduates undertaking the Anatomy and Physiology course | Quasiexperimental study design    | To compare student performance among undergraduate students with access to a virtual learning platform to students without access to the platform. | Api: Not applicable as the study evaluated multiple laboratories. There is no explicit description of the laboratories. Virtual tool: Web-based cardiovascular system platform Mode of delivery: Blended Synchronous. Students with access to the virtual platform took advantage of the virtual platform, with EG showing higher performance than CG. | Students with access to the virtual platform took advantage of the virtual platform, with EG showing higher performance than CG. Students agreed that both the animal-using laboratories and the virtual physiology laboratories were effective to support technical learning, however, they differed in terms of the skill sets that they developed. |

Table 3.—Continued
Table 3. — Continued

| Reference          | Participants                  | Study Design                        | Research Aims                                                                 | Laboratory                                      | Key Findings                                                                 |
|--------------------|--------------------------------|--------------------------------------|-------------------------------------------------------------------------------|------------------------------------------------|------------------------------------------------------------------------------|
| Modica et al.       | 227 medical undergraduates    | Quasiexperimental study design       | To assess the effectiveness of the experimental curriculum on teaching students the musculoskeletal exam as compared with a traditional curriculum. | Aim: To enhance students' conceptual understanding and technical skills in musculoskeletal examination. | No significant differences between the two groups for the examination. Students viewed the PFCs (virtual laboratory component) as the most useful in supporting their conceptual learning. The additional resources offered through the blended model of virtual physiology laboratory made the blended model the preferred mode of delivery to traditional laboratories for students. |
|                    |                                | EG, n = 120, completed the experimental curriculum, including web-based musculoskeletal tutorial, pathophysiology-focused cases, and facilitator preparation. |                                                                                   | Virtual tool: A web-based curriculum, including musculoskeletal tutorial, pathophysiology-focused cases, and facilitator preparation. Mode of delivery: Blended Asynchronous (web-based tool) and synchronous (practice session) |                                                                 |
|                    |                                | CG, n = 107, completed a traditional musculoskeletal curriculum. Multiple-choice examination and student perception survey were carried out. |                                                                                   |                                                                                   |                                                                 |

EG, experimental group; CG, control group. *Mode of delivery refers to the form of learning undertaken by the studies in offering the virtual instruments. † Only low-fidelity simulator is applicable, as the high-fidelity simulator (nasco smartscope simulator) is a mannikin simulator that can only be used in physical laboratory settings. ‡ Blended is not included as group B experienced a transition from having only animal laboratories to having only virtual laboratories.

Synthesis of Results

The effectiveness of the virtual physiology laboratories for students' conceptual, motivational, and technical learning was defined as an achievement of the laboratory in nine type percent. The effectiveness of the virtual physiology laboratories for students' conceptual learning is summarized in Fig. 2. Ninety percent of the studies conducted the physiology laboratories using blended models and 45.5% of the studies adopting both the blended and the standalone virtual models adopted both the blended and the standalone virtual models of physiology laboratories. Are virtual laboratories effective for technical learning? Six out of the eleven studies investigated whether the virtual physiology laboratories were effective for technical learning. Five out of the six studies comparing blended models with in-person laboratories (41), and one study comparing both virtual and blended models with an in-person model (40), found that virtual laboratories supported conceptual learning. The additional resources offered through the blended model of virtual physiology laboratory made the blended model the preferred mode of delivery to traditional laboratories for students. Are virtual laboratories effective for motivational learning? All three studies investigating whether the virtual physiology laboratories were effective for motivational learning supported this, with three of these studies using a blended model for the virtual laboratories. Are virtual laboratories effective for conceptual learning? Three out of the eleven studies investigated whether the virtual physiology laboratories were effective for conceptual learning. All three studies (100.0%) supported this, with most of these studies using a blended model for the virtual laboratories. A breakdown of the effectiveness of virtual laboratories for conceptual learning is summarized in Fig. 2. Ninety percent of the studies investigated whether the virtual physiology laboratories were effective for technical learning. Five out of the six studies comparing blended models with in-person laboratories (41), and one study comparing both virtual and blended models with an in-person model (40), found that virtual laboratories supported conceptual learning. The additional resources offered through the blended model of virtual physiology laboratory made the blended model the preferred mode of delivery to traditional laboratories for students. Are virtual laboratories effective for technical learning? Six out of the eleven studies investigated whether the virtual physiology laboratories were effective for technical learning. Five out of the six studies comparing blended models with in-person laboratories (41), and one study comparing both virtual and blended models with an in-person model (40), found that virtual laboratories supported conceptual learning. The additional resources offered through the blended model of virtual physiology laboratory made the blended model the preferred mode of delivery to traditional laboratories for students. Are virtual laboratories effective for conceptual learning? Three out of the eleven studies investigated whether the virtual physiology laboratories were effective for conceptual learning. All three studies (100.0%) supported this, with most of these studies using a blended model for the virtual laboratories.
DISCUSSION

This study used a systematic review to determine if virtual physiology laboratories are effective for student learning. Analysis of the 11 included studies found that virtual laboratories supported conceptual, motivational or technical learning. However, there were limitations with the mode of delivery (63.6% of the studies used blended models) and the number of studies, which will be covered in this DISCUSSION.

Conceptual Learning

The majority of the studies that investigated conceptual learning showed that the virtual physiology laboratories were effective for students’ conceptual learning. This was tested using quantitative analysis of test scores, with the “pretest-posttest” design (13, 24, 38, 39) or the “posttest-only” design (7, 16, 25, 40–42). The pretest-posttest design allows the measurement of baseline values to be used as references to show the significance of changes before and after the intervention (43), whereas the posttest-only design measures only the performance of the participants after the intervention without baseline values, which can induce biases and errors (44). Therefore, the studies that utilized the pretest-posttest design showed a higher level of validity and reliability as compared with the studies that utilized the posttest-only design.

It is likely that the mode of delivery of the virtual laboratory will also have impacts on conceptual learning. Three studies compared physiology laboratories using standalone virtual models and in-person models (see Fig. 2). Specifically, the study by Dobson (40) compared an in-person exercise physiology laboratory with a virtual simulation through a Virtual Physiology of Exercise Laboratory (VPEL) program. The study by Durand et al. (41) compared students’ perceptions and conceptual understanding of physiology laboratories that used animals and simulated physiology laboratories. In conclusion, both studies reported that the standalone virtual physiology laboratories were as effective (for conceptual learning) as in-person physiology laboratories. On the other hand, the study by Motz et al. (38), which compared a computer-based neurophysiology laboratory with an in-person neurophysiology laboratory, had inconsistent experimental and control groups, which invalidated the comparison between them. Nonetheless, their analysis of the pre- and posttest scores of only the experimental (standalone virtual laboratory) group revealed that the standalone virtual laboratory was effective for students’ conceptual learning.

The blended model of virtual physiology laboratories was found to exhibit an equal, if not a higher, level of effectiveness for students’ conceptual learning than in-person laboratories. Among the five studies that compared physiology laboratories in the blended and in-person contexts, the studies by Chen et al. (24) Dantas and Kemm (7), Elmer et al. (25), and Gopal et al. (13) compared multiple physiology laboratories in their virtual and blended models, whereas the study by Modica et al. (42) compared only the muscle physiology.
laboratories with virtual and blended models. The studies by Chen et al. (24) and Gopal et al. (13) reported that their blended models of physiology laboratories were more effective than the in-person laboratories for conceptual learning, while the studies by Dantas and Kemm (7), Elmer et al. (25), and Modica et al. (42) showed that their blended models of physiology laboratories were as effective as the in-person physiology laboratories.

Interestingly, the study by Wang et al. (16), which compared both a standalone virtual model and a blended model of neurophysiology laboratories (virtual and in-person) against an in-person neurophysiology laboratory involving live tissue dissections, showed that while the blended model was as effective as the traditional laboratory, the standalone virtual laboratory was less effective than the in-person laboratory. Hence, it showed that a blended laboratory curriculum with both virtual and physical components exhibits greater educational benefits than a standalone virtual model.

The study by Mutlu et al. (39) did not support the efficacy of virtual physiology laboratories for students’ conceptual learning. They compared an advanced technology (mannikin simulator), which was embedded into an in-person laboratory, with a less technically advanced virtual simulation (computer simulation) of the physiology laboratory. Therefore, although the virtual simulation was deemed less effective than the mannikin simulator, this supported the fact that students’ conceptual learning is effectively supported by implementing advanced technology into in-person physiology laboratories.

Therefore, for this systematic review, virtual physiology laboratories were found to be effective for students’ conceptual learning, with evidence that the blended model is an equal, if not a more beneficial, mode of physiology laboratory, to virtual laboratories.

Motivational Learning

Most of the studies that investigated motivational learning reported that virtual physiology laboratories supported this, with the majority of the students expressing that the virtual forms of physiology laboratories motivated them to actively explore physiology concepts. They evaluated motivational learning using qualitative analysis (subjective feedback) based on questionnaires and surveys for students (16, 24, 25, 37, 41) and educators (24).

The mode of delivery of the virtual laboratory could impact motivational learning. The study by Durand et al. reported that the standalone virtual physiology laboratories were effective for motivational learning by stimulating students’ interest, however, this was to a lesser extent than the in-person laboratories. The study by Moreno-Ger et al. (37), which investigated the educational impacts of a prelaboratory virtual simulation for a cardiovascular physiology laboratory, reported that the blended model of physiology laboratory was effective in stimulating students’ motivational learning by presenting the concepts in forms that are easier to understand. Similarly, Chen et al. (24) and Elmer et al. (25) reported that the blended models of physiology laboratories were more effective for students’ motivational learning than the in-person models because the virtual physiology laboratories sparked students’ interest to perform the in-person experiments.

The study by Wang et al. (16) reported that the blended model of physiology laboratory was more effective for students’ motivational learning as compared with the in-person and the standalone virtual laboratories. This is likely because the prelaboratory simulations of the experiments allow the students to practice the procedures, therefore enhancing their confidence to perform the experiments in-person.

On the other hand, the study by Dantas and Kemm (7), which tested a web-based laboratory program that included a virtual physiology laboratory as a subset of the program, showed that the program did not stimulate students’ motivational learning. Due to the design of the program, it could not be determined if the perception of poor motivational learning was attributed to the virtual laboratory component or the entire program. A potential confounding variable is the occurrence of technological advances, including the improvement of user interface and usability of interactive programs, which occurred between 2008 and the other relatively recent studies. As the improvements are expected to enhance the user experience of virtual instruments, motivational learning is likely to be stimulated. Therefore, the absence of motivational learning in this study may be attributed to the lack of realism of the student interface of the virtual physiology laboratory.

As shown, although the studies largely supported that virtual physiology laboratories are effective for students’ motivational learning, the small number of studies assessed and the subjective evaluations lowered the reliability of these evaluations. Thus the evidence is inconclusive in showing that virtual physiology laboratories are effective for students’ motivational learning.

Technical Learning

Compared with the conceptual and motivational learning, a smaller number of studies investigated the effectiveness of virtual physiology laboratories for students’ technical learning. The three studies that evaluated technical learning reported that virtual physiology laboratories were effective for students’ technical learning. This was examined using quantitative and qualitative analysis, based on a mixture of test scores (16, 37) and student perceptions (16, 41). It is worth noting that all studies investigated only the efficacy of virtual physiology laboratories for procedural skills, without testing other aspects of technical learning, such as research skills.

Based on students’ scores, Moreno-Ger et al. reported that the blended model of virtual laboratory was more effective than the in-person laboratory for students’ technical learning (37). After practicing the experiment using the prelaboratory simulation, the students showed a higher level of comprehension of the experimental procedures, and a higher level of precision when performing the experiment in-person. Whereas Wang et al. (16) reported that the blended model, the standalone virtual model, and the in-person model were equally effective for students’ technical learning as all students showed an equal understanding of the experimental equipment and procedures. Based on student perceptions, Durand et al. (41) reported...
that although the virtual physiology laboratories were as effective as the in-person laboratories in developing students’ technical skills, they differed in terms of the skill sets that they developed. For example, the in-person laboratories developed students’ “hands-on skills” through tactile experiences and the virtual physiology laboratories developed students’ information technology skills through manipulating online software to actively explore physiology concepts. Although all analyzed studies showed that virtual physiology laboratories are effective for students’ technical learning, the small number of studies assessed lowers the generalizability of this finding, therefore, this finding is inconclusive and requires further validations.

**Alignment with Past Reviews**

Prior systematic reviews have investigated the effectiveness of virtual laboratories across science disciplines such as chemistry and biology (11, 12, 27, 45, 46). A common difficulty for the past and the current reviews is the lack of standard terminology to define the variety of unconventional laboratory classes that are delivered virtually. For instance, a “remote laboratory” has been defined in some studies as being delivered entirely online while others defined it as a hands-on laboratory that can be performed outside of the conventional laboratory environments (27). Moreover, the lack of consistency in the evaluative parameters of student learning further complicates the research findings. For example, studies that supported virtual laboratories over in-person laboratories were largely based on academic performances (assessed using test scores), whereas studies that supported in-person laboratories over virtual laboratories were largely based on student perceptions (assessed using surveys) (45).

One similarity across all systematic reviews of virtual laboratories is that the included studies focused predominantly on the effectiveness of virtual laboratories for students’ conceptual learning, less on motivational learning, and rarely on technical learning (12, 46). Furthermore, the aforementioned reviews found that conceptual learning was examined largely through test scores using the pretest-posttest design or the posttest-only design (quantitative analysis) and, to a lesser extent, through student perceptions (qualitative analysis). Since quantitative analyses are generally associated with fewer experimental biases and a higher power as compared with qualitative analyses, studies that utilized quantitative analyses show higher reliability. As prior reviews largely assessed studies that utilized quantitative analyses, their finding that virtual laboratories are effective for students’ conceptual learning robustly supports this conclusion of the current review.

In comparison, investigations of students’ motivational learning associated with virtual laboratories showed conflicting results due to the large variety of virtual laboratories. According to Lewis (11), the design (e.g., asynchronous/synchronous delivery) of the virtual laboratory plays a significant role in determining student engagement and motivation to explore science concepts. Moreover, Faulconer and Gruss (27) agreed that the virtual laboratories must be designed to be highly engaging to maintain student motivation, an aspect which is absent in many existing virtual laboratory models. In this current systematic review, the designs of virtual physiology laboratories that were shown to motivate students’ learning were simulations (i.e., virtual experiments) and video demonstrations (16, 24, 25, 37, 41), whereas a web-based laboratory course was not associated with student motivational learning (7). Another factor to consider is the synchronicity of the virtual laboratories; however, we found no clear association between synchronicity and student learning. In summary, the design of virtual physiology laboratories is a potential confounding variable that may have caused the variations in students’ perceptions of motivational learning. As a result, despite the largely supportive view of virtual laboratories in effectively stimulating students’ motivational learning, further research into the educational impacts of specific designs of virtual laboratories is needed.

**Limitations and Future Implications**

The majority of the studies (7 out of 11) for this review investigated virtual physiology laboratories in the blended delivery mode, making it difficult to attribute student learning solely to the virtual component of the laboratory. Furthermore, none of the studies specifically investigated the effectiveness of virtual physiology laboratories for the development of students’ research skills, a crucial element of physiology laboratories. We also found that the use of the search term “undergraduate” led to the exclusion of a couple of papers. However, initial searches with the term “student” were too broad and led to thousands of papers from the pre-university, or school, sector. Future studies could broaden the coverage of evaluations by using synonyms of “undergraduates” such as “college” and “higher education.” Lastly, all of the studies did not provide detailed descriptions of the laboratory models used, which hindered the potential to provide recommendations for virtual laboratory models. Therefore, future studies should also include detailed reports of laboratory models.

Future studies are warranted to investigate whether specific designs of virtual laboratories support student learning. Further studies are also required to answer the question as to whether virtual laboratories are effective for research skills development. These should use valid and reliable research strategies, such as having large, randomized, and demographically consistent experimental and control groups, with participants assessed by quantitative analyses, high response rates, and rigorous data analysis.

**Conclusions**

This systematic review has critically evaluated the existing literature to determine the effectiveness of virtual physiology laboratories for student learning (conceptual, motivational, and technical). The findings of this study suggest that virtual physiology laboratories are effective for students’ conceptual learning. However, it is inconclusive as to whether virtual physiology laboratories are effective for motivational and technical learning, with further studies required to investigate this.

**DISCLOSURES**

No conflicts of interest, financial or otherwise, are declared by the authors.
41. Durand M. D T, Restini CBA, Wolff AC, Faria M, Couto LB, Bestetti RB. Students’ perception of animal or virtual laboratory in physiology practical classes in PBL medical hybrid curriculum. *Adv Physiol Educ* 43: 451–457, 2019. doi:10.1152/advan.00005.2019.

42. Modica RF, Thundiyil JG, Chou C, Diab M, Von Scheven E. Teaching musculoskeletal physical diagnosis using a web-based tutorial and pathophysiology-focused cases. *Med Educ Online* 14: 4508, 2009. doi:10.3885/meo.2009.Res00301.

43. Salkind NJ. *Encyclopedia of Research Design*. Thousand Oaks, CA: SAGE Publications, Inc, 2010, vols. 1-0.

44. Frey B. *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*. Thousand Oaks, CA: SAGE Publications, Inc., 2018, vols. 1-4.

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

46. Mirçik ÖK, Saka AZ. Evaluation of research related to virtual physics laboratory applications. *Can J Physics* 96: 740–744, 2018. doi:10.1139/cjp-2017-0747.