August 31, 2021

Effectiveness of synchronous online learning in an occupational context: two rapid reviews

Institute for Work & Health
Research Excellence
Safe Work
Healthy Workers

If you have questions about this report, please contact us at:

Institute for Work & Health
400 University Avenue, Suite 1800
Toronto, Ontario M5G 1S5

info@iwh.on.ca
www.iwh.on.ca

Institute for Work & Health

Effectiveness of synchronous online learning in an occupational context: two rapid reviews

Authors: Lynda S. Robson, Emma Irvin, Kathy Padkapayeva, Momtaz Begum, Melanie Zukowski

Date: August 31, 2021

Table of Contents

Acknowledgements

The authors thank Cameron Mustard (IWH) and Mike Parent (Workplace Safety North) for identifying, early on in the COVID pandemic, the issue pursued in this research. We thank Morgane Le Pouesard for helping screen research abstracts, Joanna Liu for library support and Maggie Tiong for developing and implementing the bibliographic search strategy.

Several organizations provided practitioner expert input on the research questions, bibliographic search terms and feedback on draft findings and the Executive Summary: Public Services Health and Safety Association, Workplace Health and Safety Centre, Workplace Safety North, Workplace Safety and Prevention Services, and the Training and Awareness Branch of the Ministry of Labour, Training and Skills Development.

The Institute for Work & Health is an independent, not-for-profit organization that operates with core funding from the Province of Ontario.

List of Abbreviations

AEL, asynchronous e-learning

COVID, coronavirus disease

F2F, face-to-face

IWH, Institute for Work & Health

OHS, occupational health and safety

SOL, synchronous online learning

List of Tables

List of Figures

Executive Summary

Introduction

The COVID pandemic and the accompanying restrictions on gatherings have precipitated a shift in Ontario and other jurisdictions away from in-person classroom delivery of training and education, including occupational health and safety (OHS) instruction. A common response to the pandemic has been the adoption of a 'virtual classroom' modality, referred to in this report as synchronous online learning (SOL), in which a learner accesses instruction through a videoconferencing platform at the same time it is delivered elsewhere by an instructor. As society plans for a post-COVID world, the question arises of whether there should be a complete return to the training modalities being used pre-COVID or whether the use of newly adopted modalities should be continued. A key consideration in this decision should be the research evidence on SOL in comparison with in-person learning (also known as face-to-face learning).

To help provide this evidence, the Institute for Work & Health (IWH) undertook two related reviews of the research literature. Of primary interest were evaluative studies of SOL to prevent adverse OHS outcomes. However, it was anticipated at the outset that such research may be difficult to find. The scope of the research included in the reviews was broadened, therefore, to training and education of an occupational (or career preparation) nature. It was assumed that findings on that type of training could reasonably be generalized to the more specific category of OHS training. Both reviews had the following **primary research question**:

- What is the effectiveness of synchronous online learning (SOL) for occupational purposes, relative to face-to-face (F2F) learning or to other e-learning?

We report here on:

1. Rapid review of systematic reviews published 2010 to 2020
2. Rapid review of primary research published in 2020.

Methods

Methods for both reviews were based on interim guidance by the Cochrane Group on rapid reviews. In each review, seven bibliographic databases were searched,

including those specializing in health sciences, education and psychology. Peer-reviewed publications meeting the following inclusion criteria were sought:

- Adult learners
- SOL of a professional/occupational/career preparation nature
- Has comparison of SOL with F2F learning or other e-learning
- Measured effects of training on knowledge, practice, or final outcome, including injuries

Steps of title and abstract screening, article screening, methodological quality assessment and data extraction were carried out by two independent reviewers.

Results and evidence synthesis from review 1

Three systematic reviews (two meta-analyses and one narrative review) were found in the research literature, that met the inclusion criteria and were of sufficient methodological quality for data extraction. The designs of studies included in the reviews were strong, mostly randomized controlled trials. Study subjects were mostly health care professionals and students of health professions (especially nursing and medicine). The reviews were consistent in providing evidence of there being no difference between SOL and F2F learning in their effects on knowledge/skills.

One of the meta-analyses, using studies of health care practitioners, mental health professionals and undergraduate students, provided evidence that the effect of SOL, relative to asynchronous e-learning (AEL), on knowledge/skill outcomes is either none or small (favouring SOL), depending on the outcome. The available evidence on the effect of SOL compared to AEL is, therefore, considered to be sparse and somewhat inconsistent.

Results and evidence synthesis from review 2

From four medium-quality studies of undergraduate/graduate students comparing SOL and F2F learning in their effects on post-training knowledge, mixed results were found (no difference and SOL more effective). There were also serious validity concerns regarding the studies.

From three medium-quality studies of undergraduate/graduate students comparing SOL and AEL in their effects on (pre-)post-training knowledge, no difference was

found. Validity and reproducibility of the results was a concern for two of these studies.

Discussion

We have contributed to the research evidence base by conducting two rapid reviews focusing on comparative effectiveness studies of SOL and F2F learning in an occupational context, which have potential application to OHS training. The ‘virtual classroom’ modality has proliferated in response to the COVID pandemic but has seldom been the focus of reviews.

The literature to address the research question was relatively sparse. This means the evidence base is not yet robust and meta-analyses cannot yet effectively explore factors explaining differences across studies in their results.

Significant gaps were found in the research literature, which need to be addressed by future research. The available evidence is derived primarily from individuals in or graduated from programs at the undergraduate level or higher. With one exception, there were no studies of people educated in community colleges, vocational colleges, or apprenticeship programs, or of people with high school preparation only. It remains to be seen whether the results found here are generalizable to a wider range of individuals, with different types of education preparation and academic abilities. Similarly, the types of training studied in the reviews was restricted in their range. In the first review, training was on clinical topics, with some training aiming for the acquisition of new clinical skills, whereas in the second review, subjects were mostly academic in nature (e.g., mathematics, pedagogy). It remains to be seen whether the results found here are generalizable to a wider range of training types, including hands-on, practical types. Finally, none of the publications included in the two rapid reviews were concerned with OHS training.

Our finding of SOL and F2F learning being of similar effectiveness is consistent with a broader literature lying outside the scope of our review, that compares either distance learning or a broad range of e-learning with F2F learning. The findings in our reviews of similar effects for SOL and AEL are also consistent with the broader literature.

Conclusions

Based on the relatively sparse evidence base examined in this review, we conclude that, generally, when all else is held equal, SOL and traditional F2F learning for occupational or career preparation purposes are similarly effective for learners at the undergraduate level or higher. This is also the case for SOL and AEL, though the evidence base is even sparser.

Research evidence is lacking on the relative effectiveness of SOL for i) other types of learners ii) hands-on, practical types of training and iii) OHS training.

Implications for practice

Given the nature of available research evidence, one can reasonably assume learning achievement using F2F and SOL will be similar when delivering content of a theoretical nature to learners at an undergraduate-level, if appropriate consideration has been given to instructional principles in both modalities. Such similarity of achievement cannot yet be assumed when delivering content of a more hands-on practical nature, or when delivering content to adult learners with less educational preparation, since research is lacking in these areas. Although OHS training has not been specifically studied in the research reviewed here, we expect these comments would apply to it too.

Introduction

Background and rationale

The COVID pandemic and the accompanying restrictions on gatherings have precipitated a shift in Ontario and other jurisdictions away from in-person delivery of training and education, including occupational health and safety (OHS) instruction. A common response to the pandemic has been the adoption of a 'virtual classroom' modality, referred to in this report as synchronous online learning (SOL), in which a learner accesses instruction through a videoconferencing platform at the same time it is delivered by an instructor.

The pandemic hastened ongoing societal trends towards greater use of e-learning,[1] including both synchronous and asynchronous online learning, instead of face-to-face (F2F) learning – trends driven by the potential to decrease costs and enhance accessibility. However, as society plans for a post-COVID world, the question arises of whether there should be a return to training modalities being used pre-COVID or whether the use of newly adopted modalities should be continued. On the one hand, e-learning offers potential cost savings to all stakeholders; on the other hand, concerns persist among decision-makers that e-learning may be less effective than the traditional F2F learning. A key consideration in this decision should be the research evidence on the effectiveness of e-learning in comparison with F2F learning.

The two reviews presented below sought research evidence on the relative effectiveness of one type of e-learning, SOL. This is defined as Internet-mediated distance education/training delivered in 'real-time' by an instructor. SOL often includes opportunities for the learner to interact with the instructor through audio or chat function. A SOL course delivery may include some asynchronous elements too, such as online learning modules, resources, and asynchronous chats.

[1] We define e-learning "an approach to teaching and learning which uses electronic media or devices to provide access to training or education", adapting a longer definition by Sangra et al. (2012).

Of primary interest in the reviews were evaluative studies of SOL to prevent adverse OHS outcomes. However, it was anticipated at the outset that such research would be difficult to find, especially that involving a comparison with F2F learning. The scope of the research included in the reviews was therefore broadened to include all training and education of an occupational or career preparation nature. It was assumed that research on that type of training could be reasonably generalized to the more specific category of OHS training.

Report structure: two rapid reviews

The two related reviews reported on here are:

1. Rapid review of systematic reviews published 2010 to 2020
2. Rapid review of primary research published in 2020.

Separate Methods and Results are presented for each review, followed by a joint Discussion and Conclusion section.

Rapid reviews are a newer form of knowledge synthesis derived from systematic reviews (Moher et al., 2015). Rapid review methodology streamlines some review processes, preserving a systematic and rigorous approach, but yielding a synthesized evidence product in a more timely manner (Garrity et al., 2021).

The decision to conduct the first review as a review of systematic reviews, rather than a review of primary studies, was driven by the large number of records identified in preliminary literature searches and the need for timely results. However, since only three systematic reviews on SOL were ultimately identified and none involved OHS training, a decision was subsequently made to undertake the second rapid review of primary research. It was limited to 2020 to capture studies not represented within the systematic reviews included in first rapid review, including those published in response to the COVID pandemic.

Research question

Both rapid reviews had the following primary research question:

- What is the effectiveness of synchronous online learning (SOL) for occupational purposes, relative to face-to-face (F2F) learning or other e-learning?

Methods – Review 1

Methods for both reviews were based on interim guidance by the Cochrane Group on rapid reviews (Garrity et al., 2021). Steps of screening, data extraction, and quality assessment were managed through DistillerSR software (Evidence Partners Limited).

Stakeholder consultation

At the outset of the first review, the IWH research team contacted publicly-funded organizations in Ontario involved with OHS training to assess interest in the research question and in e-learning more generally. Respondents gave feedback and input on a proposed set of eligibility criteria and set of search terms; planned methods were modified accordingly. In a second round of consultation, organizations provided feedback on the Executive Summary, following a slide presentation of the results, to ensure that messages were clear. Five organizations, listed in the Acknowledgements section participated in one or both rounds of consultation.

Search strategy and information sources – review 1

Seven bibliographic databases were searched for systematic reviews, including those specializing in health sciences, education and psychology: Cochrane Library, Education Source, Embase, ERIC, MEDLINE(OVID), PsycINFO, Scopus. The search strategy, developed with a professional librarian, identified all titles and abstracts that had i) one or more training/education/e-learning search terms, ii) one or more training outcome terms, and iii) a systematic/scoping review term. The outcome terms were both OHS-specific (e.g. injury) and generic (e.g., knowledge, behaviour, practice). The search, conducted November 2, 2020, also used another set of terms to exclude children and was restricted to English language publications from 2010-2020. An illustrative detailed search strategy is in Appendix A.

In addition to the bibliographic searches, the following additional information sources were screened: reference lists of included articles, reference lists of relevant overviews identified through the bibliographic database searches, and systematic/scoping reviews identified in Review 2. In addition, we followed up on study protocols identified through the database searches to determine if they had been subsequently published.

Study selection process and eligibility criteria – review 1

Each of the titles and abstract records identified through the above searches were independently screened for eligibility by two reviewers, following a pilot training period involving senior reviewers. For any title and abstract meeting our inclusion criteria, full papers were retrieved for independent screening by two reviewers. At both steps, conflicts in reviewer assessment were resolved through discussion.

Eligibility criteria were refined as the team became more familiar with the literature, with decisions guided by the aim of being able to apply the results of the review to OHS instruction. For this reason, included studies were required to involve training or education delivered in a professional, occupational or career preparation context. The final version of eligibility criteria is in Table 1 and the screening forms are in Appendix B.

[Link to Table 1]

Data extraction – review 1

Information about methodology, population, intervention, comparisons and outcomes was independently extracted by two reviewers for all eligible reviews (data extraction form in Appendix C). Data for outcomes were extracted if they met the eligibility criteria. Differences were resolved through discussion and a consensus version was developed. Reviewers relied on the information in the article and accompanying online supplementary materials only.

Quality assessment – review 1

The AMSTAR 2 measurement tool (Shea et al., 2017) was used to assess the methodological quality of each eligible review. Two reviewers carried out their independent assessment using the form in Appendix D and resolved any discrepancies through discussion. Reviewers did not screen, appraise, or extract data on any study they authored or co-authored.

Evidence synthesis – review 1

A narrative synthesis of the evidence was developed by one researcher (LR) and was reviewed critically by two other researchers (EI, KP).

Results – Review 1

Study selection – review 1

The study selection process is summarized in Figure 1. From 1,653 unique records of publications, identified through the searches of bibliographic databases and other information sources, 236 were identified as potentially relevant to the review.[2] Full papers were retrieved for 234 of these and screened for eligibility. Full papers subsequently excluded consisted of 161 for which e-learning was not a primary focus and 70 for which an e-learning was the main focus. Of this latter group, 67 were excluded (see Appendix E for references) for one of the following reasons: e-learning in studies did not involve SOL; e-learning in studies might have involved SOL but it was unclear; or reviews contained only single studies of SOL. There were three exclusions because of the latter reason (Jackson et al., 2018; Khurshid et al., 2020; Richmond et al., 2017). This left three eligible publications (Chipps et al. 2012; Gegenfurtner and Ebner, 2019; He et al., 2020), listed in Appendix F.

[Link to Figure 1]

Study characteristics – review 1

The three eligible reviews consisted of two meta-analyses and one narrative systematic review. All three focused on SOL as the intervention of interest, referring to it as videoconference-based education, synchronous webinar and synchronous distance education. The number of primary studies available in each review to address this overview's research question was modest, ranging from five to nine.

The designs of studies included in the reviews were strong. The two meta-analyses included only randomized controlled trials in their primary analyses. [One of these two included non-randomized controlled trials too in a sensitivity analysis.] The narrative review included randomized controlled trials, randomized crossover trials

[2] Screening and eligibility criteria at the title and abstract stage were inclusive of all types of remote e-learning. At the full paper screening stage, eligibility criteria were narrowed to focus on SOL as shown in Table 1.

and one single case alternating treatment study. The average number of participants per study, calculated for each review, was modest, ranging from 50 to 120.

Study subjects were primarily from nursing or medicine, either including both professionals and students (Chipps et al., 2012; Gegenfurtner and Ebner, 2019) or restricted to students (He et al., 2020). Studies included in Chipps et al. (2012) were published between 2003 and 2008, while those in the other two reviews were published in the following decade. There was overlap between the two later studies, with five studies included in both reviews. Eligible outcomes synthesized by the reviews included only immediate knowledge/skill achievement. (The Chipps et al. (2012) review also presented results the following measures from a single study: confidence-post, knowledge-follow-up, skill-follow-up.) The characteristics of the eligible reviews are presented in Table 2.

[Link to Table 2]

Study methodological quality – review 1

The assessments of the methodological quality of the reviews with the AMSTAR 2 instrument (Table 3) indicate some limitations in the review methodologies. Of particular note was the lack of consideration given to the methodological quality of the individual studies (see AMSTAR 2 items #9, #12, #13) especially with the two meta-analyses. Although He et al. (2020) carried out two different types of methodological quality assessments, and noted no studies had a high risk of bias (p. 6), results were not considered in the conduct of the meta-analysis or interpretation of findings. Gegenfurtner and Ebner (2019) did not assess methodological quality nor discuss it. Chipps et al. (2012) gave attention to methodological quality but their assessment method was not fully described. Criteria #2, #7, #10, #16 were also not well met across the reviews but these were considered by the research team to be of relatively low importance in the present context of non-commercial educational interventions.

[Link to Table 3]

There were strengths found in the reviews too. In every review, eligibility criteria were clear and sensible. Four to 11 databases were searched with an extensive set of terms; and were supplemented by searches of reference lists of included articles or a

search for articles citing the included articles. Finally, publication bias was examined in the two meta-analyses.

Results found in individual studies – review 1

Effects of SOL versus F2F on knowledge and/or skills

Figure 2 summarizes the five results available in the two meta-analyses (Gegenfurtner and Ebner, 2019; He et al., 2020) addressing the research question. Each result is a pooled estimate of 'effect size'[3] derived from pooling the effect size estimates calculated from the knowledge/skill outcomes in individual studies. In this case, effect size expresses the difference in effects between SOL and F2F learning, such that a positive value means SOL is more effective than F2F. The five pooled effect size estimates, each derived from three to seven studies, range from 0.02 to 0.16, which is very small to small by conventional standards (Cohen, 1988). Furthermore, from a statistical point of view, these values are not statistically significant, since their 95% confidence intervals encompass the value of zero.

[Link to Figure 2]

All five studies included in the narrative review (Chipps et al., 2012) had findings for post-training knowledge. Four showed no statistically significant difference between SOL and F2F in effects and one showed a highly significant effect favouring SOL ($p < 0.004$). Effect size information was not available in, nor calculable from the data provided in the review. Amongst the five studies, there were additional potentially eligible outcome measures reported (knowledge integration, follow-up knowledge, skills, follow-up skills, confidence) but they were each reported for only a single study and are thus not included here.

Additional statistical findings

Statistical heterogeneity was reported on in only the He et al. (2020) study. For the post-training-only measures of knowledge and skill, heterogeneity was low (I^2 = 0-

[3] Effect size is a unitless standardized measure, allowing comparison across studies, even when different measures (e.g. different knowledge tests) are used. A common convention (Cohen, 1988) considers effect sizes of 0.2 as small, 0.5 as medium, and 0.8 as large.

18%), whereas for the pre-post training knowledge measures, I^2 = 55%, which is considered moderate (Higgins et al., 2021).

There was no evidence of publication bias, either graphically or statistically, when examined in the two meta-analyses.

Effects of SOL versus asynchronous e-learning (AEL) on knowledge and/or skills

Only one of the meta-analyses (Gegenfurtner and Ebner, 2019) compared the effects of SOL and asynchronous e-learning (AEL). Two outcomes were examined in samples of community health aides/practitioners, mental health professionals and undergraduate students: knowledge/skills post-only and knowledge/skills pre-post. The pooled effect size estimate[4] for knowledge/skills post-only, based on five effects from three studies, was small, positive, and statistically significant, favouring SOL: 0.217 (0.045 – 0.389); whereas for knowledge/skills pre-post, based on four effects from two studies, was close to zero and not significant: 0.044 (-0.130 – 0.218).

Evidence synthesis – review 1

Two meta-analyses and one narrative review of studies, involving mostly health care professionals and students of health professions (especially nursing and medicine), were consistent in providing evidence of there being no difference between SOL and F2F learning in their effects on knowledge/skills. Since the reviews included a modest number of primary studies and the reviews had some methodological limitations, despite being mostly randomized controlled trials, the available evidence is not yet definitive, and its robustness to variation in learners, curriculum, and setting is unknown.

One meta-analysis based on health care practitioners, mental health professionals and undergraduate students provided evidence that SOL relative to AEL on knowledge/skill outcomes has either no effect (pre-post outcomes from two studies) or a small effect, favouring SOL (post-only outcomes from three studies). The available evidence was therefore considered to be sparse and somewhat inconsistent.

[4] Cohen’s d corrected for sampling error

Methods – Review 2

Search strategy and information sources – review 2

Seven bibliographic databases were searched for primary studies, including those focused on health sciences, education and psychology: Cochrane Library, Education Source, Embase, ERIC, MEDLINE(OVID), PsycINFO, Scopus. The search strategy, developed with a professional librarian, identified all titles and abstracts that had i) one or more training/education/e-learning search terms, ii) one or more training outcome terms. The outcome terms were both OHS-specific (e.g., injury) and generic (e.g., knowledge, behaviour, practice). The search also used another set of terms to exclude the education of children and was restricted to English language publications from 2020. The search was conducted on December 16, 2020. An illustrative detailed search strategy is in Appendix G. In addition to the bibliographic searches, the reference lists of included articles were screened but yielded no additional references.

Study selection process and eligibility criteria – review 2

Following a pilot training period involving senior reviewers, all titles and abstract records were screened for eligibility by a single reviewer (MB), while 20% were screened by a second reviewer. The review team met weekly to resolve emergent issues. For any title and abstract classified as potentially relevant, full papers were retrieved for independent screening by two reviewers. Conflicts in reviewer assessment were resolved through discussion. Eligibility criteria are in Table 4 and the screening forms are in Appendix H.

[Link to Table 4]

Quality assessment – review 2

Methodological quality assessment was carried out using a 23-item instrument developed by the IWH for use with a variety of study designs. Assessment items are primarily concerned with aspects of internal validity, including intervention allocation, similarity of comparison groups, compliance, contamination, co-intervention, outcome measurement, attrition and control of confounding. See Appendix I for a copy of the instrument. Two reviewers carried out their independent assessment and resolved discrepancies through discussion.

An overall score was derived by the assessment as follows: A = sum of (item weight x item response score); B = sum of (item weight x maximum response score); overall score (%) = A/B x 100. Item weight was 1, 2 or 3 and decided upon by consensus among LR, EI, and MB. Items concerned with internal validity and relevant to the body of literature were prioritized.

Data extraction – review 2

Information about methodology, population, intervention, comparisons and outcomes was independently extracted by two reviewers for all eligible articles (data extraction form in Appendix J). Data for outcomes were extracted from tables if they met the eligibility criteria. Differences were resolved through discussion and a consensus version was developed. Reviewers relied on the information in the article and accompanying online supplementary materials. In the case of Root and Rehfeldt (2021), additional material from the same set of studies, reported on in a PhD dissertation (Root, 2019), was included in data extraction. While reviewers did not routinely request additional information from authors, exceptions were made to avoid creating errors in the presentation of material in this report. Corresponding authors for Dahlstrom-Hakki et al. (2020) and Morice et al. (2020) were contacted to seek clarification on the specifics of the training delivery method. Dr. Root was contacted because of discrepancies within the statistical information presented in Root (2019). He recommended we conduct our own statistical tests on the study data. We therefore conducted paired t-tests on the data provided in Tables 1 and 4 in Root (2019) and Figure 3 in Root and Rehfeldt (2021). Reviewers did not screen, appraise, or extract data on any study they authored or co-authored.

Evidence synthesis – review 2

A narrative synthesis was developed by one researcher (LR) and was reviewed critically by two other researchers (EI, KP).

Results – Review 2

Study selection – review 2

The study selection process is summarized in Figure 3. From 6,571 unique records of publications identified through the searches of bibliographic databases, 270 were identified as potentially relevant to the review. Full papers could be retrieved for 265 of these and were screened for eligibility. Ten of these were found to be eligible (references listed in Appendix K).

[Link to Figure 3]

Study characteristics – review 2

Table 5 summarizes the characteristics of the 10 studies included in review 2. Study designs were generally weak; only one was a randomized controlled trial. The weakest designs -- nonconcurrent cohort and an observational before-after study – were found among the four studies that had arisen as a result of the COVID pandemic. Sample sizes in some studies were very small, with less than thirty per study group in Dela Cruz et al. (2020), Guo (2020), Mullen (2020) and Root and Rehfeldt (2021).

Study subjects were most often undergraduate or graduate students in health care (3 studies) education (2 studies) or other fields (4 studies); in one study they were clinicians. There was wide variation in the total SOL time involved in the training/education studied, from 1 hr to 240 hrs. Most often, studies compared SOL to F2F learning, but there were some comparisons to AEL or other types of SOL. The outcomes studied were only those, collected immediately post-training, sometimes with a baseline measure having been collected: (pre-)post knowledge and (pre-)post self-efficacy.

[Link to Table 5]

Study methodological quality – review 2

Results from assessing the methodological quality of the 10 studies are shown in Table 5 and Figure 4, with detailed results available in Appendix L. Scores out of 100 ranged from 10 to 69. Six were classified to be of medium quality, i.e., receiving a score of 50% or higher. From these studies, results data were extracted and are

presented below. The remaining four studies received a score below 50% (Carroll et al., 2021; Dela Cruz et al., 2020; Guo et al., 2020; Morice et al., 2020) and did not progress further in the review.

[Link to Figure 4]

Results from individual studies – review 2

The training interventions in the six studies classified as medium-quality, shown in Table 6, were varied in topic: statistics, global health, education, behavior analysis and suicide-counseling. They varied in duration too, from one to 45 hours, as well as in the nature of the training intervention.

[Link to Table 6]

Effects of SOL versus F2F on post-training knowledge – review 2

Four studies contributed evidence of effectiveness of SOL relative to F2F learning, based on post-training knowledge measurement. Lee et al. (2020), Mullen (2020) and Root and Rehfeldt (2021) compared entire courses delivered in either SOL or F2F modalities; course length varied from 11 hours taught over 2 days to 45 hours taught over 15 weeks. Norton (2020) compared workshops delivered in either SOL or F2F modalities, though in both cases lectures were delivered as SOL and online resources were also provided.

Findings were mixed, as shown in Table 7. In the cases of Lee et al. (2020), the second exam of Norton (2020), and Root and Rehfeldt (2021), SOL was found to be more effective than F2F learning. In contrast, there was no difference found in the cases of Mullen (2020) and the first exam of Norton (2020).

[Link to Table 7]

There are significant validity concerns with these studies. For Lee et al. (2020) and Root and Rehfeldt (2021), we note the inherently weak study designs. In particular, data on SOL in Lee et al. (2020) was collected during the COVID pandemic, whereas data on F2F learning was collected pre-pandemic – two very different contexts for the medical students under study. The validity of the study design used by Root and Rehfeldt (2021) depends on the quizzes being compared across modalities being of equal difficulty (i.e., the five quizzes assessing learning in the alternating five weeks of SOL instruction needed to be of equal difficulty as the five quizzes assessing

learning in the alternating five weeks of F2F instruction.) However, this equivalency was not demonstrated by the authors. Norton (2020) did not adjust in his analysis for the differing characteristics of the compared groups, especially their undergraduate preparation in math, which he cited as explaining why the two groups differed in their results on the second exam (requiring higher math understanding), but did not differ on first exam. Finally, though Mullen (2020) used a quantitative measure to assess knowledge, namely the number of pages as a proxy for quality of submitted assignments, her determination of the equivalency of the total number of pages was not statistical in nature (considered them equal though there was a 3% difference). As well, the total sample size of 14 was extremely small, leading to concerns about the reproducibility of results.

Effects of SOL versus AEL on knowledge outcomes – review 2

Three studies from two publications contributed evidence on the effectiveness of SOL, relative to AEL, based on knowledge measured either post-training only or both pre- and post-training (Table 8). The nature of the interventions compared were quite varied, with Dahlstrom-Hakki et al. (2020) comparing asynchronous versus synchronous discussion sessions, Root and Rehfeldt (2021) comparing synchronous online lectures to self-paced learning modules in Experiment 2, and the same authors comparing synchronous lectures plus asynchronous online discussion (Intervention 1) to self-paced modules plus synchronous chat (Intervention 2) in Experiment 3. No statistically significant difference in the effects of the training modalities was seen.

[Link to Table 8]

We note that sample sizes were very small in the Root and Rehfeldt (2021) experiments, so the question arises as to whether no difference was seen because statistical power was low. This may be the case, but it is also true that effects were small, with absolute differences in the between-modality effects of 0.2 and 1.6 percentage points out of 100, respectively, for Experiments 2 and 3 (see Appendix N). However, standard deviations were large (9.6 and 8.6 percentage points, respectively), suggesting this finding may not be readily reproducible in a new sample of students. As discussed in the above section, uncertainty about the robustness of the result also arises from the study design.

Evidence synthesis – review 2

From four studies of undergraduate/graduate students comparing SOL and F2F learning in their effects on post-training knowledge, mixed results were found (no difference and SOL more effective). There were also serious validity concerns.

From three studies of undergraduate/graduate students comparing SOL and AEL in their effects on (pre-)post-training knowledge, no difference was found, but the validity and reproducibility of these results was a concern in two of the studies because of study design and sample size.

Discussion

Two rapid reviews were conducted to assess the effectiveness of SOL, relative to F2F learning and to other e-learning: 1) a rapid review of systematic reviews published 2010-2020 and 2) a rapid review of primary studies published in 2020. The first review identified three eligible systematic reviews from which we have extracted evidence. The second review identified 10 eligible primary studies, of which six were assessed to be of sufficient methodological quality from which to extract evidence.

We have contributed to the research evidence base by focusing on training delivered in an occupational context, which has potential application to OHS training.

Evidence on the SOL versus F2F learning comparison

The three systematic reviews included in the first rapid review were consistent in presenting evidence of SOL and F2F being similar in their effects on immediate post-training knowledge/skill outcomes. Small, non-significant effects favouring SOL were seen in the two meta-analyses. In the narrative systematic review, there was one study favouring SOL and four evidencing no difference. Evidence in the second rapid review, based on four primary studies was mixed (either no difference or favouring SOL), though it carries less weight than the evidence of the first review, because of greater methodological concerns.

While the balance of evidence available to date indicates no difference between the SOL and F2F learning, this could change with the addition of more studies and more precise meta-analytic estimates. Studies of the SOL versus F2F learning contrast are not yet plentiful. The pooled estimates in the included systematic reviews were based on no more than seven individual studies. However, a sensitivity analysis conducted by He et al. (2020) on post-training knowledge, including not only seven randomized controlled trials, but also 21 non-randomized two-group studies, derived an estimate indicating no difference too (effect size = -0.002 (-0.11 – 0.10)).

Evidence on the SOL versus AEL comparison

Evidence comparing SOL to AEL was sparser than for the previous contrast. In the first rapid review, one meta-analytic study, using data from three primary studies, provided evidence of either a small positive effect favouring SOL or no effect, depending on the choice of knowledge/skill outcome – post-only and pre-post,

respectively. Results from the second review, based on three primary studies reported in two publications, indicated no difference in effects between modalities, but methodological quality was a concern. On balance, we consider that the very sparse evidence available suggests SOL and AEL are similar in their effect.

Limitations in the methodology used in the two rapid reviews

The research team chose to address the research question with rapid review methods so that results could be produced in the most timely way. The resulting choices of selecting only English-language articles and minimizing search strategies other than bibliographic searches may have led to the omission of relevant articles. On the other hand, the scope of the bibliographic search was broad, including databases abstracting publications in health sciences, psychology and education.

Another way of containing the scope of the work was to restrict the eligible publication dates to 2010 and later, with the rationale that videoconferencing technologies and SOL have become more advanced and widespread only in recent years. Retrieved articles did identify some earlier reviews, discussed below, but they validated our decision. Earlier reviews included older technologies (e.g. Bernard et al., 2004) or grouped SOL and AEL modalities together (Cook et al., 2008).

Expediency also led to the decision to review systematic reviews rather than primary studies in the first rapid review, which covered 2010-2020. It is possible that primary studies of the effectiveness of OHS training in SOL modality were published prior to 2020, but we did not identify them through these reviews. However, the complete absence of such OHS-related studies among those included in the second rapid review, which screened primary studies released in the first 10 months of 2020, suggests that relevant OHS-related primary studies published before 2020 would likely have been few, if any.

Limited breadth of study populations and types of training

The subjects of the systematic reviews eligible for inclusion in the first rapid review were primarily students of or professionals in health care fields, especially nursing and medicine. In the second review, they were mostly undergraduate and graduate students in health, education and other fields. The available evidence is therefore derived primarily from individuals with high educational attainment. Except for one study of community health aides/practitioners, included in one of the systematic

reviews, there were no studies of people educated in community colleges, vocational colleges or apprenticeship programs, or of people with high school preparation only. It therefore remains to be seen whether the results found here are generalizable to a wider range of individuals, with different types of education preparation and academic abilities.

The available evidence is similarly restricted with regards to the content and aims of training. In the first review, training was on clinical topics, with some aiming for the acquisition of new specific clinical skills, whereas in the second review, subjects were mostly academic in nature (e.g., mathematics, pedagogy). It remains to be seen whether the results found here are generalizable to a wider range of training types, especially hands-on, practical training.

Lack of publications concerned with OHS

None of the publications included in the two rapid reviews were concerned with OHS training. We conducted audits of the excluded studies to determine whether any eligible studies had been inadvertently excluded but did not find any. It revealed that some OHS-related publications had been identified through the library search process, but they had been appropriately excluded at the abstract or full-paper screening stage. For instance, the first rapid review identified only one review focused OHS and e-learning (Gao et al., 2019), but excluded it because SOL studies were not included. We note that only one of the studies in the Gao et al. (2019) review involved a comparison of a type of e-learning to F2F learning: Sacks et al. (2013) compared construction safety training delivered in virtual reality format and traditional F2F learning. The second rapid review identified 12 English-language primary studies related to OHS at the screening stage (references listed in Appendix O) but all were excluded because they did not involve SOL. Eight of the articles were concerned with virtual reality, two with serious games, one with other asynchronous e-learning and one with a cyber-physical postural training. Two of the 12 were randomized controlled trials. One compared virtual reality to lecture-based safety training in construction (Nykänen et al., 2020) and the other compared asynchronous e-learning modules on PPE use to reading guidelines on PPE use in emergency health care (Suppan et al., 2020).

The lack of OHS-related publications included in the two rapid reviews raises the question of whether the reviews' findings would be applicable to OHS. Favouring

applicability is that articles were restricted to adults undergoing career preparation or continued learning. However, training studied in the reviews tended to be academic in nature, although some aimed for the acquisition of clinical skills. Results of the rapid reviews are therefore likely more applicable to OHS training that is theory-oriented rather than practice-oriented.

Other limitations in the literature included in the rapid reviews

The literature available to address the study's research question was relatively sparse. There were only three eligible systematic reviews available to the first rapid review, with a modest number of primary studies included in each. The number of eligible studies available to the second rapid review was also modest. This means the evidence base is not yet robust and meta-analyses cannot effectively explore factors explaining differences across studies in their results.

The literature included in the rapid reviews was also limited in the scope of outcomes studied, focusing on immediate learning outcomes, mostly knowledge/skill achievement. It appears that follow up studies measuring the transfer of learning to the work setting are seldom conducted.

Methodological quality was found to be weak among the studies eligible for inclusion in the second rapid review. Six of 10 were classified as medium quality, four of low quality, and none of high quality. Low quality was especially associated with studies of emergency remote teaching developed in response to the COVID pandemic.

Assessment of the methodological quality of the three reviews included in the first rapid review also revealed limitations, but concern is less with these. Mitigating considerations were that relatively thorough searches had been conducted in each of the three reviews and the studies included in these reviews were almost all randomized controlled trials. As well, the two reviews that investigated methodological quality (Chipps et al., 2012; He et al., 2020) revealed no major concerns. Importantly, both meta-analyses found no suggestion of publication bias.[5]

[5] Publication bias arises when statistically non-significant results in one particular direction of relative effect are not published. It is detected through visual inspection of graphs and/or statistical tests of all results.

We note that our attention to the issue of methodological quality is warranted. It has been found to explain substantial portions of variation in effect sizes in meta-analyses of training studies (Bernard et al., 2004; 2009; Cook et al., 2008)

Relation to other research literature

We have made a contribution to the research evidence base in focusing on comparative effectiveness studies of SOL and F2F learning in an occupational context. The 'virtual classroom' modality has proliferated in response to the COVID pandemic but has seldom been the focus of reviews, as it was here.

Our finding of SOL and F2F learning being of similar effectiveness is consistent with a broader literature, lying outside the scope of our review. For example, in a large meta-analysis (Bernard, 2004) comparing distance education to classroom learning in both children and adults, in which SOL delivery included not only videoconferencing but also TV/radio broadcast, a very small negative effect on knowledge outcomes, favouring F2F learning was found: effect size (95% CI) = -0.10 (-0.15, -0.06). Three other research studies have looked at Internet-based instruction in comparison with F2F learning, without separating SOL and AEL modalities. Means et al. (2013) reported a very small, non-significant positive effect (effect size = 0.05; p = 0.46), when 27 effects from studies of adults and children were meta-analyzed. Cook et al. (2008) focused on Internet-based learning for health care professionals, estimating a small advantage to acquiring knowledge (effect size = 0.21 (0.07 to 0.34)) based on 48 interventions) but not for acquiring skills (effect size = 0.06 (–0.33 to 0.44)) based on 11 interventions). Richmond et al. (2017) sought to update the work and restrict study design eligibility to randomized controlled trials. They synthesized two to four studies at a time, looking at knowledge, skill or behaviours as outcomes with either lectures or workshops as the F2F comparison. The six pooled effect size estimates ranged from -0.25 to 0.22, but in no case were statistically significant.

The findings in our reviews of similar effects between SOL and AEL are also consistent with the literature. Bernard et al.'s (2004) meta-analytic estimate for asynchronous distance education relative to classroom learning was 0.05 (0.03 – 0.08), i.e., 0.15 effect size units greater than their corresponding estimate for synchronous distance education relative to classroom learning.

Some of the outcomes examined in the He et al. (2020) meta-analytic study were found to have moderate heterogeneity, meaning the relative effectiveness of SOL may be dependent on characteristics of learners, training intervention, comparator intervention, outcome measurement, etc. Such heterogeneity has been found in the broader literature too (Bernard et al., 2004; Cook et al., 2008). A larger evidence base is needed before the basis of the heterogeneity can be well understood. Studies were too few in the Gegenfurtner and Ebner (2019) and He et al. (2020) reviews, though exploratory analyses were undertaken. It is worth noting that investigations by Bernard et al. (2004) showed that variation in study results for SOL vs F2F was most dependent on the methodological quality of studies, followed by pedagogical variables, and then the media-related variables. Accordingly, they recommended that attention to quality course design should take precedence over attention to the characteristics of media. This is echoed by the recent messages of academics in distance education, who maintain that while teaching online and face-to-face are different, the fundamentals of good quality education remain the same, requiring learner-centred, high-quality instructional design (Fawns et al., 2020; Hodges et al., 2020). They note that emergency remote education, developed too quickly in response to the COVID pandemic, might not meet these requirements.

Future research

Future research should address the research gaps identified here. Comparative research, assessing differences in the effectiveness of different training modalities, needs to be conducted with workers with less educational preparation than a university/college degree. It needs to be understood too through research how the relative effectiveness might vary with how theoretical/practical the training is. This potential moderator has not yet considered in the meta-analyses included in our rapid review, nor in the literature reviewed above. Another aspect needing more attention is longer term effects, especially learning transfer to the workplace. Finally, for decision-makers in OHS, there is a need for research on the effectiveness of different modalities in OHS-related training.

Conclusions

Based on the relatively sparse evidence base examined in this review, we conclude that, generally, when all else is held equal, SOL and traditional F2F learning for occupational or career preparation purposes are similarly effective for learners at the undergraduate level or higher. This is also the case for SOL and AEL, though the evidence base is even sparser.

Research evidence is lacking on the relative effectiveness of SOL for i) other types of learners ii) more practical types of training and iii) OHS training.

Implications for practice

Given the nature of available research evidence reviewed here, one can reasonably assume learning achievement using F2F and SOL will be similar when delivering content of a theoretical nature to learners at an undergraduate-level or above, when appropriate consideration has been given to instructional principles in both modalities. Caution should be taken with this assumption when delivering content of a more practical nature and when delivering content to adult learners with less educational preparation, since research is lacking in these areas. Although OHS training has not been specifically studied in the research reviewed here, we expect these comments would apply to it too.

References

Bernard, R.M., Abrami, P.C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., Wallet, P.A., Fiset, M., Huang, B. 2004. How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. Review of Educational Research 74(3):379-439.

Bernard, R.M., Abrami, P.C., Borokhovski, E., Wade, A., Tamim, R.M., Surkes, M.A., Bethel, E.C. 2009. A meta-analysis of three types of interaction treatments in distance education. Review of Educational Research 79(3), 1243-1289. doi:10.3102/0034654309333844

Carroll, T., Mooney, C., Horowitz, R. 2021. Re-ACT: Remote Advanced Communication Training in a time of crisis. Journal of Pain and Symptom Management 61, 364-368. doi:10.1016/j.jpainsymman.2020.08.013

Chipps, J., Brysiewicz, P., Mars, M. 2012. A systematic review of the effectiveness of videoconference-based tele-education for medical and nursing education. Worldviews on Evidence-Based Nursing 9, 78-87. doi:10.1111/j.1741-6787.2012.00241.x

Cohen, J. 1988. Statistical Power Analysis for the Behavioral Sciences (2nd Ed). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. 2008. Internet-Based Learning in the Health Professions: A Meta-analysis. JAMA. 300(10):1181–1196. doi:10.1001/jama.300.10.1181

Dahlstrom-Hakki, I., Alstad, Z., Banerjee, M. 2020. Comparing synchronous and asynchronous online discussions for students with disabilities: the impact of social presence. Computers & Education 150, 103842. doi:10.1016/j.compedu.2020.103842

Dela Cruz, A.M., Alick, S., Das, R., Brenner, A. 2020. Same material, different formats: comparing in-person and distance learning in undergraduate medical education. Academic Psychiatry 14, 659-663. doi:10.1007/s40596-020-01333-7Dahl-strom

Fantom, N., Serajuddin, U. 2016. The World Bank's classification of countries by income. Policy Research Working Paper No. 7528. World Bank, Washington, DC. © World Bank. https://openknowledge.worldbank.org/handle/10986/23628 License: CC BY 3.0 IGO.

Fawns, T., Jones, D., Aitken, G. 2020. Challenging assumptions about “moving online” in response to COVID-19, and some practical advice. MedEdPublish 9(1), 83. doi: 10.15694/mep.2020.000083.1

Gao, Y., Gonzalez, V.A., Yiu, T.W. 2019. The effectiveness of traditional tools and computer-aided technologies for health and safety training in the construction sector:

a systematic review. Computers & Education 138, 101-115. doi:10.1016/j.compedu.2019.05.003

Guo, S. 2020. Synchronous versus asynchronous online teaching of physics during the COVID-19 pandemic. Physics Education 55, 1-9. doi:10.1088/1361-6552/aba1c5

Garritty, C., Gartlehner, G., Nussbaumer-Streit, B., King, V.J., Hamel, C., Kamel, C., Affengruber, L., Stevens, A. 2021. Cochrane Rapid Reviews Methods Group offers evidence-informed guidance to conduct rapid reviews. Journal of Clinical Epidemiology 130, 13-22. doi: 10.1016/j.jclinepi.2020.10.007.

Gegenfurtner, A., Ebner, C. 2019. Webinars in higher education and professional training: a meta-analysis and systematic review of randomized controlled trials. Educational Research Review 28, 100293. doi:10.1016/j.edurev.2019.100293

He, L., Yang, N., Xu, L., Ping, F., Li, W., Sun, Q., Li, Y., Zhu, H., Zhang, H. 2020. Synchronous distance education vs traditional education for health science students: a systematic review and meta-analysis. Medical Education 55, 293-308. doi:10.1111/medu.14364

Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021). Cochrane, 2021. Available from www.training.cochrane.org/handbook

Hodges, C., Moore, S., Lockee, B., Trust, T., Bond, A. 2020. The difference between emergency remote teaching and online learning. EDUCAUSE Review. Available from https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning. Accessed August 27, 2021.

Jackson, C.B., Quetsch, L.B., Brabson, L.A., Herschell, A.D. 2018. Web-based training methods for behavioral health providers: a systematic review. Administration and Policy in Mental Health 45(4), 587–610. Doi:10.1007/s10488-018-0847-0

Khurshid, Z., De Brún, A., Moore, G., McAuliffe, E. 2020. Virtual adaptation of traditional healthcare quality improvement training in response to COVID-19: a rapid narrative review. Hum Resources for Health 18, 81. Doi: 10.1186/s12960-020-00527-2

Kirkpatrick, D. L. 1994. Evaluating training programs: the four levels. San Francisco: Berrett-Koehler.

Lee, S.J., Park, J., Lee, Y.J., Lee, S., Kim, W.H., Yoon, H.B. 2020. The feasibility and satisfaction of an online global health education course at a single medical school: a retrospective study. Korean Journal of Medical Education 32, 307-315. doi:10.3946/kjme.2020.178

Means, B., Toyama, Y., Murphy, R., Baki, M. 2013. The effectiveness of online and blended learning: a meta-analysis of the empirical literature. Teachers College Record 115(3), 1–47.

Moher, D., Stewart, L., Shekelle, P. 2015. All in the Family: systematic reviews, rapid reviews, scoping reviews, realist reviews, and more

Morice, A., Jablon, E., Delevaque, C., Khonsari, R.H., Picard, A., Kadlub, N. 2020. Virtual versus traditional classroom on facial traumatology learning: evaluation of medical student's knowledge acquisition and satisfaction. Journal of Stomatology, Oral and Maxillofacial Surgery 121, 642-645. doi:10.1016/j.jormas.2020.03.001

Mullen, C.A. 2020. Does modality matter? A comparison of aspiring leaders' learning online and face-to-face. Journal of Further and Higher Education 44, 670-688. doi:10.1080/0309877X.2019.1576859

Norton, S. 2020. Australian primary mathematics teacher preparation: on-campus or online? Who? Why? So What? Mathematics Teacher Education & Development 22, 91-114.

Nykänen, M., Puro, V. Tiikkaja, M., Kannisto, H., Lantto, E., Simpura, F., Uusitalo, J., Lukander K., Räsänen, T. Heikkilä, T., Teperi, A.-M. 2020. Implementing and evaluating novel safety training methods for construction sector workers: results of a randomized controlled trial. Journal of Safety Research 75, 205-221, DOI: 10.1016/j.jsr.2020.09.015.

Richmond, H., Copsey, B., Hall, A.M., Davies, D., Lamb, S.E. 2017. A systematic review and meta-analysis of online versus alternative methods for training licensed health care professionals to deliver clinical interventions. BMC Medical Education 17(1), 227. doi: 10.1186/s12909-017-1047-4.

Root W.B. 2019. The synthesis of programmed instruction and online education: towards a modern-day teaching machine. Dissertation Submitted in Partial Fulfillment of the Requirements for the Doctor of Philosophy Degree. Department of Rehabilitation in the Graduate School, Southern Illinois University Carbondale. 13901649.

Root, W.B., Rehfeldt, R.A. 2021. Towards a modern-day teaching machine: the synthesis of programmed instruction and online education. The Psychological Record 71, 85–94. https://doi.org/10.1007/s40732-020-00415-0

Sacks, R., Perlman, A., Barak, R. 2013. Construction safety training using immersive virtual reality. Construction Management & Economics 31(9): 1005-1017.

Sangrà, A., Vlachopoulos, D., Cabrera, N. 2012. Building an inclusive definition of e-learning: an approach to the conceptual framework. The International Review of Research in Open and Distributed Learning 13(2), 145-159. Doi: 10.19173/irrodl.v13i2.1161

Shea, B.J., Reeves, B.C., Wells, G., Thuku, M., Hamel, C., Moran, J., Moher, D., Tugwell, P., Welch, V., Kristjansson, E., Henry, D.A. 2017. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 358, j4008. doi:10.1136/bmj.j4008

Suppan, L., Abbas, M., Stuby, L., Cottet, P., Larribau, R., Golay, E., Iten, A., Harbarth, S., Gartner, B., Suppan, M. 2020. Effect of an E-learning module on personal protective equipment proficiency among prehospital personnel: Web-based randomized controlled trial. Journal of Medical Internet Research 22(8), e21265. DOI: 10.2196/21265.

Tables

Table 1: Eligibility criteria for reviews – review 1

	INCLUDE	EXCLUDE
Population	• adult learner, including apprentices and post-secondary students • country classified as Very High by the 2014 UN Human Development Index[1]	• children • countries with UN HDI classification lower than Very High
Intervention	• 'virtual classroom'/SOL • professional/occupational/career preparation training including learning clinical guidelines	• Mentoring (only) interventions • MOOCs • training to acquire language • e-learning for non-occupational health outcome • e-learning which teaches a motor skill, e.g., surgical training, ball-playing because it is not readily generalizable to OHS training • e-learning that cannot be definitively identified as SOL
Comparison	• face-to-face learning, e-learning	• no training (including pre-post)
Outcomes	• knowledge, skills, attitudes targeted by the training (including self-efficacy in new skills), practices (transferred knowledge) or "final" outcomes, including injury rates and patient outcomes (i.e. Kirkpatrick levels 2, 3, or 4)[2]	• outcomes which precede knowledge gain in the causal chain, including satisfaction, engagement, attitude toward training, motivation (Kirkpatrick Level 1), etc.
Language	• English	• Non-English

Publication Type	• Reviews of primary studies or of reviews • Reviews involving a systematic search component • Peer-reviewed	• Reviews with an unspecified or non-systematic search process • Bibliographic reviews • Primary studies • Book chapters, letters, commentary, proceedings, conference abstracts, protocols • Scoping reviews

[1] Countries classified as Very High in the 2014 UN Human Development Index (Fantom and Serajuddin, 2016) include Andorra, Argentina, Australia, Austria, Bahrain, Belgium, Brunei Darussalam, Canada, Chile, Croatia, Cuba, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hong Kong SAR, China, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Rep., Kuwait, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Netherlands, New Zealand, Norway, Poland, Portugal, Qatar, Saudi Arabia, Singapore, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, United Arab Emirates, United Kingdom, United States

[2] Kirkpatrick (1994) is a widely used framework for evaluating the outcomes of training. Levels 1 to 4 respectively are satisfaction, learning, behavior and results

Back to study selection process and eligibility criteria – review 1

Table 2: Characteristics of systematic reviews of synchronous online learning effectiveness – review 1

First author (year)	Chipps (2012)	Gegenfurtner (2019)	He (2020)
Country of 1st author	South Africa	Germany	China
Study designs included in review	RCTs; randomized crossover trials; single case, alternating treatment	RCTs	RCTs
Type of evidence synthesis	Narrative	Meta-analysis	Meta-analysis
Publication dates of included studies	2003-2008	2010-2017	2010-2019
No. of studies in the review (no. addressing RQs)	5	12 (9)	7
Total no. of participants in review	251	1414	594
Occupations	Nurses and students of HC professions (nursing, medicine, other)	Primarily HC professionals (2 studies) and students of HC professions (e.g. medicine, nursing; 5 studies), as well as one study of community health aides/practitioners and one study of undergraduate students	HC students (nursing, medical, dental, physiotherapy)

First author (year)	Chipps (2012)	Gegenfurtner (2019)	He (2020)
Interventions	SOL, undergraduate, post-graduate, and in-service lectures or workshops; 3 to 24 hr total time	SOL, professional and higher education lectures or workshops; some interactive; 0.5 to 25 hr total time	SOL: Synchronous higher education, some interactive; 0.5 hr to 1.5 day total time
Comparisons	F2F	F2F, AEL, no-training	F2F

AEL, asynchronous e-learning; F2F, face-to-face learning; HC, health care; no., number; RCT, randomized controlled trial; synchronous online learning; RQ, research question.

Back to Results – Study characteristics – review 1

Table 3: Assessment of methodological quality using AMSTAR 2 – review 1

	Chipps (2012)	**Gegenfurtner (2019)**	**He (2020)**
1. Research questions and/or inclusion criteria include PICO?	Yes	Yes	Yes
2. Review methods established prior to the conduct of the review?	**No**	**No**	**No**
3. Explain selection of included study designs?	Yes	Yes	Yes
4. Comprehensive literature search strategy?	**No** (language restrictions not discussed)	Partial Yes	Partial Yes
5. Study selection performed in duplicate?	Yes	Yes	Yes
6. Data extraction performed in duplicate?	**No** (not reported)	Yes	Yes
7. Exclusions justified and provided?	**No**	**No**	**No**
8. Describe included studies in adequate detail?	Partial Yes	Partial Yes	Partial Yes
9. Use satisfactory technique for assessing the risk of bias in included individual studies?	**No**	**No**	Yes
10. Sources of funding mentioned?	**No**	**No**	**No**
11. Meta-analytic methods appropriate?	Not applicable	Yes	Yes
12. Account for risk of bias in individual studies in results?	Not applicable	**No**	**No**
13. Account for risk of bias in individual studies in interpretation?	Yes	**No**	**No**
14. Heterogeneity explained and discussed?	Yes	Yes	**No** (not discussed)
15. Publication bias investigated and discussed?	Not applicable	Yes	Yes
16. Competing interests reported upon and any existing handled appropriately	**No**	**No**	Yes

Back to Results – Study methodological quality – Review 1

Table 4. Eligibility criteria – review 2

	INCLUDE	EXCLUDE
Population	• adult learner, including apprentices and post-secondary students • country classified as Very High by the 2014 UN Human Development Index[1]	• children • countries with UN HDI classification lower than Very High
Intervention	• "virtual classroom" training a.k.a. "virtual instructor", SOL • professional/occupational/career preparation training including learning clinical guidelines • can be used remotely by a large segment of the target population (excluding, for example, expensive e-learning equipment, such as mannequins)	• Mentoring (only) interventions • MOOCs • training for acquiring language • e-learning for non-occupational health outcome • e-learning which teaches a motor skill, e.g., surgical training, ball-playing because it is not readily generalizable to OHS training • e-learning that cannot be definitively identified as 'virtual classroom'
Comparison	• face-to-face learning or other e-learning	• no training (including pre-post)
Outcomes	• Kirkpatrick Level 2,[2] such as knowledge/skill/targeted attitude (including self-efficacy in new skill); Level 3, such as behaviour/practice; Level 4 or "final outcome" for individual/organization (i.e. injury rate, patient outcome)	• Outcomes which precede knowledge gain in the causal chain (Kirkpatrick 1), including satisfaction with, engagement with, attitude toward or motivation fo r e-learning, etc.
Language	• English	• Non-English
Publication Type	• Primary study, including PhD dissertations	• Book chapters, letters, commentary, editorials, newspaper articles, proceedings, conference abstracts, protocols, Master's dissertations

[1] See Table 1 for list of countries and citation. [2] See Table 1 for explanation of Kirkpatrick levels.

Back to Study selection process and eligibility criteria – review 2

Table 5: Study characteristics and methodological quality – review 2

First Author, year	Country	Study design (* = planned due to COVID)	n total	Occupation/ type of student	SOL learning time	Eligible contrast(s) from study[1]	Outcome(s)	Quality score
Carroll, 2021	USA	Nonconcurrent cohort*	150	Interdisciplinary clinicians	45-60 mins	SOL vs F2F	Self-efficacy-post	10
Dahlstrom-Hakki, 2020	USA	Non-randomized crossover trial (6 treatment phases)	105	College students with disabilities (learning, ADHD, autism, other)	3 hrs	SOL vs AEL	Knowledge-pre-post	60
Dela Cruz, 2020	USA	Nonconcurrent cohort*	44	3rd yr (SOL) and 4th yr (F2F) medical students	240 hrs	SOL vs F2F	Self-efficacy-post	32
Elliott, 2020	USA	Randomized controlled trial	90	1st yr counseling students	I1: 1 hr I2: 1 hr	SOL(1) vs SOL(2)	Self-efficacy-pre-post	69
Guo, 2020	USA	Before-after (observational) within cohort*	9 (within cohort of 20)	2nd and 3rd yr undergraduate science students	38 hrs	SOL vs F2F	Knowledge-pre-post	44
Lee, 2020	Korea	Nonconcurrent cohort*	299	2nd yr medical students	11 hrs	SOL vs F2F	Knowledge-post	60
Morice, 2020	France	Cohort	84	5th yr medical students	3 hrs	SOL vs F2F	Knowledge-post	38
Mullen, 2020	USA	Cohort	14	2nd yr Master of Education students (K-12 teachers)	45 hrs	SOL vs F2F	Knowledge-post	62
Norton, 2020	Australia	Cohort	342	Master's students (Primary education pre-service teachers)	28 hrs	SOL vs F2F	Knowledge-post	56

First Author, year	Country	Study design (* = planned due to COVID)	n total	Occupation/ type of student	SOL learning time	Eligible contrast(s) from study[1]	Outcome(s)	Quality score
Root, 2021 (and Root, 2019)	USA	E1, E2, E3: Single case, alternating (randomized) treatments E1: 10 phases E2: 9 phases E3: 9 phases	E1:24 E2: 8 E3:10	E1: Undergraduates E2: Graduate students E3: Graduate students	E1: 9.2 hrs E2: 5 hrs E3-I1: SOL lectures: 4 hrs E3-I2 SOL chat: 5 hrs	E1: SOL vs F2F E2: SOL vs AEL E3: SOL lectures + AEL discussion (I1) vs AEL modules + SOL chat (I2)	Knowledge-post	63

Abbreviations: AEL, asynchronous e-learning; E1, experiment 1; est., estimated by reviewers from data provided in publication; F2F, face-to-face; n, total number of participants in SOL intervention and comparison groups; KSA, Kingdom of Saudi Arabia; nr, not reported; SOL, synchronous online learning

[1] Interventions/comparisons in table are restricted to those addressing the research questions and may not include all groups in a study.

[2] Supplemented with information from Root, 2019.

Back to Results – Study Characteristics – Review 2

Table 6: Characteristics of SOL and comparison interventions (medium-quality only) – review 2

Study first author, year	Learning topic	SOL intervention components key to contrast studied	No. of SOL sessions	SOL learning time (total)	SOL – period of delivery	Other components of SOL intervention[1]	Comparison condition relative to entire SOL intervention
Dahlstrom-Hakki, 2020	Introductory statistics	Instructor- prompted discussion with video, audio, chat	3	3 hrs	semester	Optional AEL module prior to each session. Full F2F instruction delivered after posttest for each session.	Same except discussion was asynchronous, available over 24 hrs.
Elliott, 2020	Suicide-specific counseling skills	I1: Role-playing I2: Q&A	I1 = 1 I2 = 1	I1: 1 hr I2: 1 hr	I1: 1 hr I2: 1 hr	Week-long AEL-dominant module, with reading, a reflective paper, and *practice with another person*	N/A, because two SOL interventions compared
Lee, 2020	Global health	Live-streamed video lectures (also recorded), group discussions, presentations, guest lectures, talk shows	1	11 hrs	2 days	N/A	F2F: same components, all F2F; lasted *14 hrs rather than 11 hrs*
Mullen, 2020	Curriculum leadership (master level)	Instructor presentation; facilitation of student assignments	15	45 hrs	15 wks	N/A	F2F: same components as SOL, all F2F

Study first author, year	Learning topic	SOL intervention components key to contrast studied	No. of SOL sessions	SOL learning time (total)	SOL – period of delivery	Other components of SOL intervention[1]	Comparison condition relative to entire SOL intervention
Norton, 2020	Mathematics curriculum	Workshops, with 2-way interaction through audio or chat; recorded	14	28 hrs	14 wks	Weekly SOL lectures (2 hr); additional videos and written material including tools	F2F: Same as SOL except workshops were F2F and *in addition, access to recordings of the SOL workshops*
Root, 2021 (and Root, 2019)	E1: Introductory applied behavior analysis E2: Graduate-level ethics course in behavior analysis E3: Graduate-level course in behavior analysis	E1: Lectures, incorporating within-lecture assessment questions, chat for questions E2: Lectures E3-I1: Lectures E3-I2: Live chat	E1: 10 E2: 5 E3-I1: 4 E3-I2: 5	E1: 9.2 hrs E2: 5 hrs E3-I1: 4 hrs E3-I2: 5 hrs	E1: 5 wks E2: 5 wks E3-I1: 4 wks E3-I2: 5 wks	E1: N/A E2: N/A E3-I1: mandatory participation in 2 asynchronous online discussion boards per week E3-I2: AEL; self-paced e-learning modules	E1: F2F: same components, though nature of F2F interaction nr. E2: AEL: 3 self-paced modules per week E3: N/A – two SOL interventions compared

AEL, asynchronous e-learning; est., estimated from data provided in publication; F2F, face-to-face; hrs, hours; I1, intervention 1; I2, intervention 2; N/A, not applicable; nr, not reported; Q&A, question and answer; SOL, synchronous online learning; wk, week

[1] Other components of SOL intervention were delivered as SOL unless otherwise indicated.

Back to Results from individual studies – review 2

Table 7: Results on the relative effectiveness of SOL versus F2F learning environments on post-training knowledge – review 2

Study 1st author, year published	Study design n total	Post-training knowledge measure	Author's conclusion about effects (p-value for between-training modality statistical test)
Lee, 2020	Nonconcurrent cohort n = 299	Final exam score (%)	SOL > F2F (p<0.005)
Mullen, 2020	Cohort n = 14	# of pages in 4 submitted assignments (proxy for quality)	SOL = F2F (nr)
Norton, 2020	Cohort n = 342	Exams on math and math pedagogy, scores out of 100 Exam 1: lower primary grades Exam 2: upper primary grades	F2F not more effective than SOL: Exam 1: SOL = F2F (p = 0.13) Exam 2: SOL > F2F (p = 0.04) Exam 2 differences attributed to pre-existing group differences, especially prior math training)
Root, 2021; E1	Single case, alternating treatment, 10 phases n = 24	% correct in each of 10 quizzes	Reviewers[1]: SOL > F2F p = 0.0003

Further details of results are found in Appendix M.

Abbreviations: #, number; E1, experiment 1; n, number of people in study; nr, not reported; sd, standard deviation.

[1] Reviewers' calculations based on raw data in Table 1 of Root (2019), with permission of Dr. Root.

Back to Effects of SOL versus F2F on post-training knowledge – review 2

Table 8: Results on the relative effectiveness of SOL versus AEL on knowledge outcomes – review 2

Study 1st author, year published	Study design n total	Outcome	Outcome measurement Statistical approach	Author's/ reviewers' conclusion about effects (p-value for between-training modality statistical test)
Dahlstrom-Hakki, 2020	Non-randomized crossover trial, 6 treatment phases n = 105	Knowledge-pre-post	Test scores from 12 tests GLMEM	Author's: SOL= AEL ($p < 0.1$)
Root, 2021; E2	Single case, alternating treatments, 9 phases n = 8	Knowledge-post	% correct in each of 9 quizzes paired t-test	Reviewers[1]: SOL = AEL ($p = 0.68$)
Root, 2021; E3	Single case, alternating treatments, 9 phases n = 10	Knowledge-post	% correct in each of 9 quizzes paired t-test	Reviewers[2]: I1 (SOL lecture + AEL discussion) = I2 (AEL modules + SOL chat) ($p = 0.56$)

Further details of results found in Appendix N.

Abbreviations: AEL, asynchronous online learning; E2 and E3, experiments 2 and 3; GLMEM, generalized linear mixed effects modeling; I1 and I2, interventions 1 and 2; sd, standard deviation; SOL, synchronous online learning.

[1] Reviewers' calculations based on raw data in Table 4 of Root (2019), with permission of Dr. Root.

[2] Reviewers' calculations based on data estimated from Fig. 3 of Root & Rehfeldt (2021), with permission of Dr. Root.

Back to Effects of SOL versus AEL on knowledge outcomes – review 2

Figures

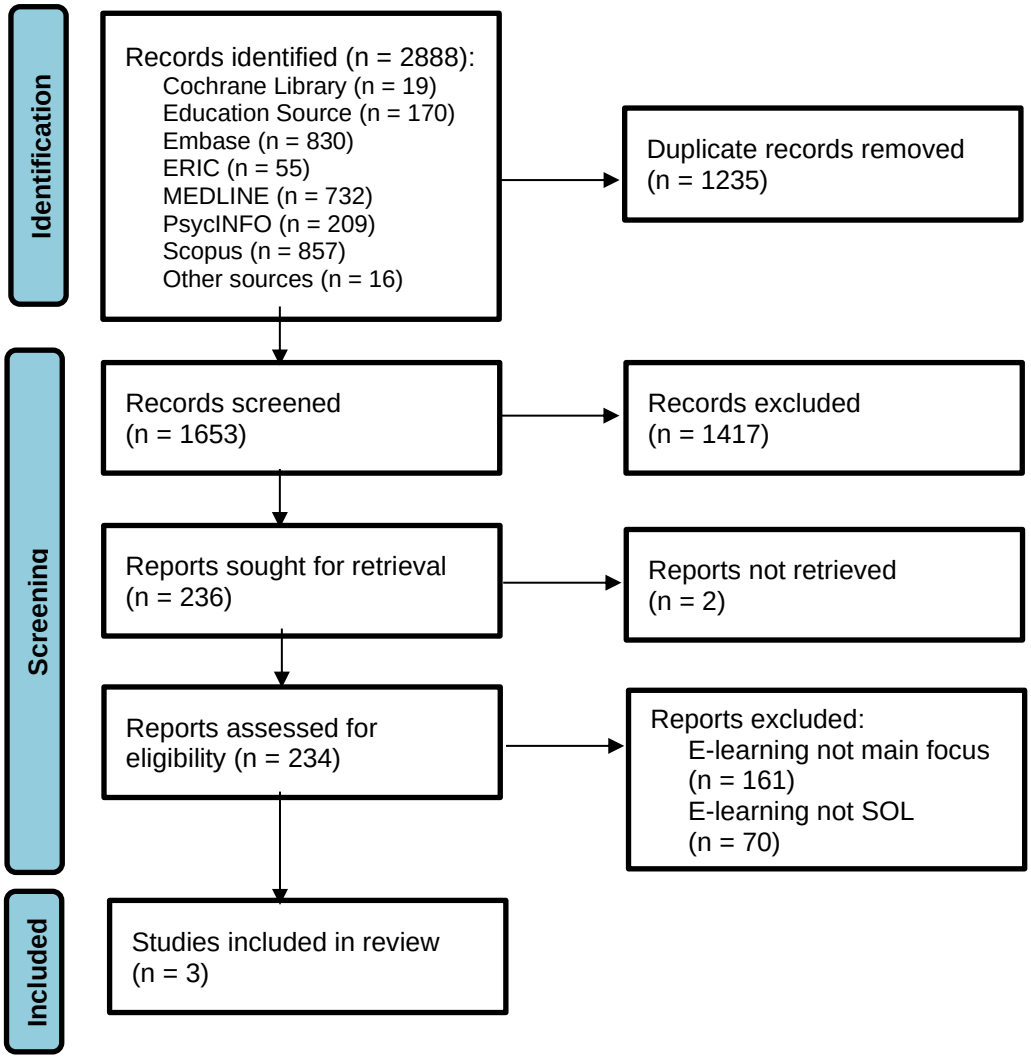

Figure 1: Summary of study identification and selection – review 1

Back to Results – Study selection – review 1

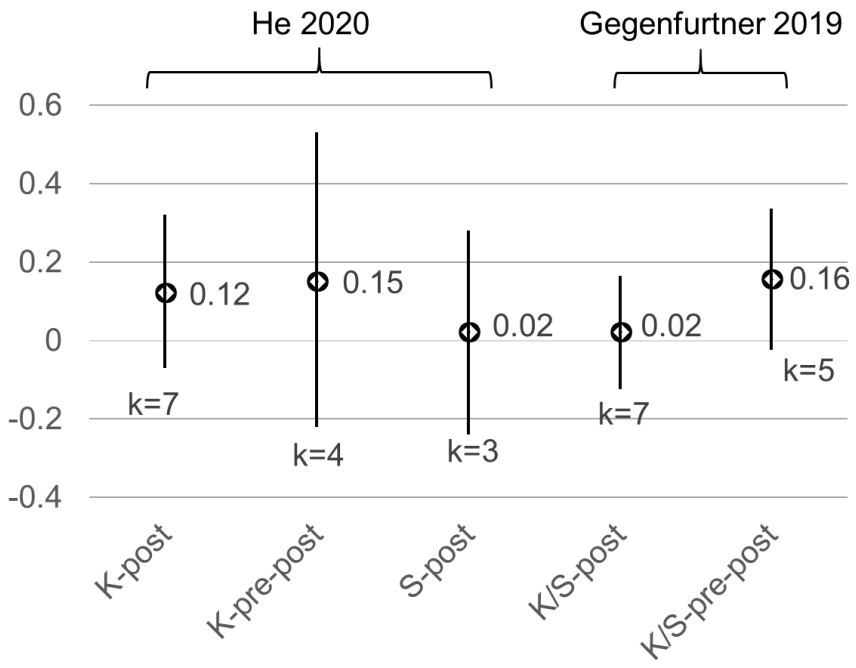

Figure 2: Summary of results from two meta-analyses on difference in effects of SOL and F2F learning – review 1

Positive values indicate SOL is more effective than F2F. In the case of He et al. (2020), pooled estimates of Hedge's g and their 95% confidence intervals are shown and k is the number of effects contributing to the pooled estimate, each from an independent study. In the case of Gegenfurtner and Ebner (2019), pooled estimates of Cohen's d corrected for sampling error and their 95% confidence intervals are shown (taken from Tables 9-10 in the source publication) and k is the number of knowledge/skill effects contributing to the pooled estimate, with some studies contributing both a knowledge and a skill effect. "Post" outcome measures are based on knowledge/skill measured post-training only and "pre-post" measures are based on them measured both pre- and post-training. Abbreviations: K, knowledge; S, skill.

Back to Results found in individual studies – review 1

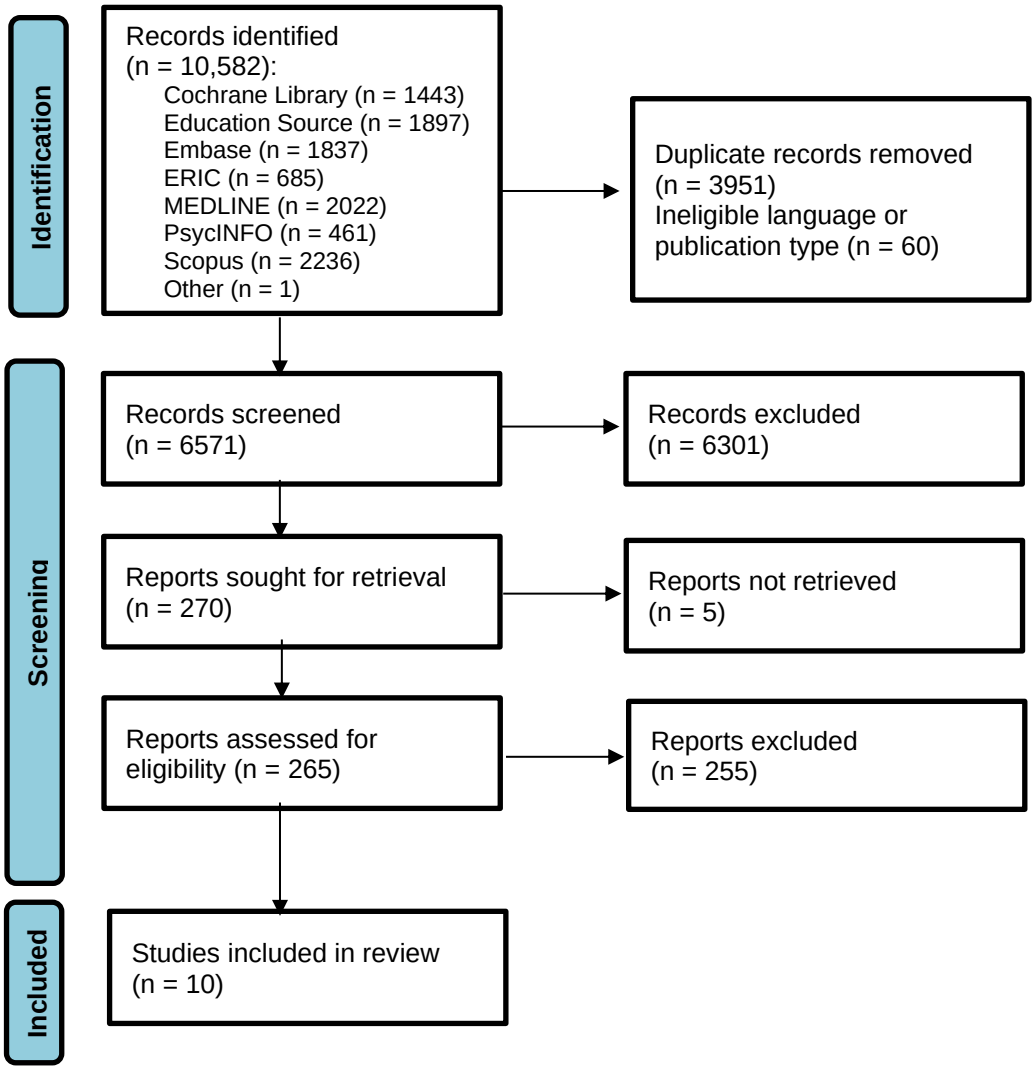

Figure 3: Summary of study identification and selection – review 2

Back to Results – Study selection – review 2

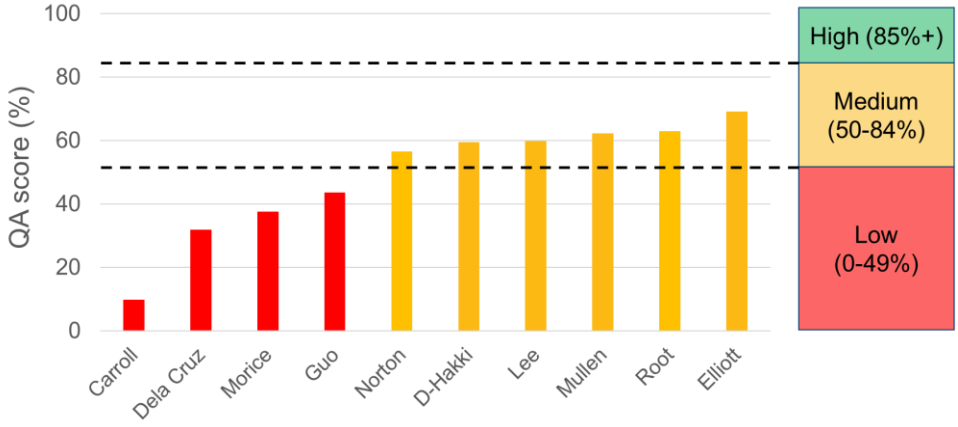

Figure 4: Distribution of study methodological quality scores – review 2

Back to Results – Study methodological quality – Review two

Appendices

Appendix A: Search strategy for Medline – review 1

2 November 2020
N=732
Database: Ovid MEDLINE: Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE® Daily and Ovid MEDLINE® <1946-Present>
Search Strategy:
---

1 exp Education, Distance/ (4249)
2 exp Computer-Assisted Instruction/ (11929)
3 (distance adj2 learning).ti,ab. (1400)
4 (distance adj2 education).ti,ab. (928)
5 (distance adj2 instruct*).ti,ab. (28)
6 (distance adj2 teach*).ti,ab. (75)
7 (distance adj2 training).ti,ab. (363)
8 e-learning.ti,ab. (2505)
9 elearning.ti,ab. (243)
10 (e-training or e-trainer$).ti,ab. (70)
11 (etraining or etrainer$).ti,ab. (2)
12 "online classroom*".ti,ab. (43)
13 (online adj2 course$).ti,ab. (1317)
14 (online adj2 education).ti,ab. (1194)
15 (online adj2 learning).ti,ab. (1879)
16 (online adj2 lecture$).ti,ab. (179)
17 (online adj2 instruct*).ti,ab. (1627)
18 (online adj2 teach*).ti,ab. (423)
19 (online adj2 training).ti,ab. (1021)
20 (remote adj2 instruct*).ti,ab. (39)
21 (remote adj2 learning).ti,ab. (98)
22 (remote adj2 teach*).ti,ab. (61)
23 (remote adj2 training).ti,ab. (131)
24 "virtual classroom*".ti,ab. (128)
25 (virtual* adj2 learning).ti,ab. (635)
26 (virtual* adj2 instruct*).ti,ab. (54)
27 (virtual* adj2 live).ti,ab. (55)
28 (virtual adj2 teach*).ti,ab. (94)
29 (virtual* adj2 train*).ti,ab. (1001)
30 "virtual reality".ti,ab. (10026)
31 "augmented reality".ti,ab. (2086)
32 "web-based training".ti,ab. (280)
33 "training simulator$".ti,ab. (303)
34 "training lab$".ti,ab. (217)

35 ((synchronous or asynchronous) adj2 learning).ti,ab. (139)
36 "spaced learning".ti,ab. (66)
37 "ubiquitous learning".ti,ab. (8)
38 u-learning.ti,ab. (4)
39 (MOOC or "massive open online course$").ti,ab. (270)
40 "information and communication technology".ti,ab. (1436)
41 ICT.ti,ab. (5503)
42 or/1-41 (42082)
43 exp child/ (1924944)
44 "K-12".ti,ab. (8250)
45 ("elementary school$" or "middle school$" or "junior high" or "high school$" or "secondary school$").ti,ab. (56618)
46 43 or 44 or 45 (1972258)
47 42 not 46 (39929)
48 accident$.ti,ab. (113731)
49 exp Accidents/ or exp Accidents, Occupational/ (192098)
50 exp Occupational Health/ (33826)
51 best practices.ti,ab. (13574)
52 behavio?r.ti,ab. (844995)
53 claim$.ti,ab. (95141)
54 Evaluation study/ (254485)
55 Health Behavior/ (50754)
56 injur*.ti,ab. (812469)
57 knowledge.ti,ab. (715562)
58 exp Occupational Diseases/ (132646)
59 exp Occupational Exposure/ (63985)
60 (occupation* adj2 exposure).ti,ab. (20772)
61 practice.ti,ab. (715327)
62 safety.ti,ab. (519062)
63 Safety Management/ (20293)
64 (work* adj2 engag*).ti,ab. (3657)
65 (work* adj2 exposure).ti,ab. (5945)
66 (work adj2 participat*).ti,ab. (1126)
67 (work* adj3 practice*).ti,ab. (12125)
68 exp Workers' Compensation/ (7524)
69 "Wounds and Injuries"/ (77525)
70 skill$.ti,ab. (202919)
71 survey$.ti,ab. (658137)
72 test$.ti,ab. (3254233)
73 score$.ti,ab. (952755)
74 grade$.ti,ab. (429656)
75 learning.ti,ab. (304359)
76 or/48-75 (7966231)
77 ("systematic review" or "scoping review").ti,ab. (173557)
78 "Systematic Review"/ (137942)

79 77 or 78 (197204)
80 47 and 76 and 79 (658)
81 47 and 76 (25994)
82 limit 81 to "reviews (maximizes specificity)" (770)
83 80 or 82 (850)
84 83 (850)
85 limit 84 to english language (833)
86 limit 85 to yr=2010-2020 (732)

Appendix B: Screening forms (Levels 1, 2, 3) – review 1

Screening form – Level 1

1. **Is the publication a review, using a systematic search method and synthesizing evidence about e-learning effectiveness?** (Yes, No, Can't Tell)
2. **Is the e-learning of an occupational/professional/career-preparing nature?** (Yes, No, Can't Tell)
3. **Are one or more eligible outcomes reported upon in a summative quantitative way: knowledge, skill, behaviour/practice, or "final outcome"?** "*Summative quantitative" means they have summarized e-learning results separately in a quantitative way by counting study results or pooling effects. "Final outcome" is what the new knowledge/skill is ultimately meant to affect (e.g. patient/hospital outcome, learner/organizational safety)*. (Yes, No, Can't Tell)
4. **Are study participants adults in a developed country? (If all countries included in review, answer YES/CAN'T TELL). Developed country defined as Very High in UN HDI[3].** (Yes, No, Can't Tell)
5. **If applicable, flag this record for possible future use.?** (Conceptual contribution, Bibliographical study, Protocol for eligible study, Key reference cited which is outside of review window, OHS training review or single study, Relevant scoping review, Other)
6. **Notes**

Screening form – Level 2 (Full paper screening)

1. **Is the publication a review, using a systematic search method and synthesizing evidence about e-learning effectiveness?** (Yes, No)
2. **Is the e-learning of an occupational/professional/career-preparing nature?** (Yes, No)
3. **Is one or more eligible outcomes reported upon in a summative quantitative way: knowledge, skill, behaviour/practice, or "final outcome"?** (Yes, No)
4. **Are study participants adults in a developed country? (If unstated or all countries, select YES).** (Yes, No)
5. **If applicable, flag this record for possible future use.** (Conceptual contribution, Bibliographical study, Protocol for eligible study, Key reference cited which is outside of review window, Relevant scoping review, Identifies primary studies about e-learning, OHS training review or study, Other)
6. **Notes**

Screening form – Level 3 (Full paper screening II)

1. **Is the publication a systematic review with the effectiveness of one or more e-learning types a major focus of the study?** (Yes (include); No, but it seems valuable to keep anyways, given our research questions (include); No, it is an overview of systematic reviews (exclude); No, (exclude).
2. **Is the evidence for virtual classroom versus face-to-face classroom or other e-learning presented separately from other evidence?** *Answer "no" if virtual classroom versus face-to-face classroom/other e-learning might be included in review but cannot be unambiguously disentangled from results, or if it is not available at all.* (Yes, No)
3. **Which type of comparisons are included?** (E-learning vs. no instruction, E-learning vs. face-to-face classroom instruction, E-learning I vs. E-learning II)
4. **Notes**

Appendix C: Data extraction form (Level 4) – review 1

Quality control:

1. **Should this article be excluded from data extraction because it is NOT relevant based on full article criteria?** (Article is NOT relevant based on title and abstract criteria and should be excluded. Reason(s) for exclusion; Article is relevant but data is not extractable (e.g., only one study with extractable data, unclear/insufficient detail) Explain further; Reference should be included)

Study identification and setting:

2. **Name of 1st author, year of publication and country.** (Name (last name); Year (YYYY); Country (of first author))

Study characteristics

3. **State the research question(s)/objective(s).**
4. a) **List the inclusion and exclusion criteria for the review using the P I C O S framework below.** (Population, Intervention, Control/Comparison Group, Outcome, Study Design)
   b) **What was the range of publication years allowed in the literature search? Include specific cutoff date.**
   c) **How many studies are included in the review pertaining to the within-scope questions?**
   d) **What was the range of publication years of the studies actually found and included in the within-scope portion of the review?**
   e) **How many participants are included in the review pertaining to the within-scope questions?**
5. **Please indicate the types of analyses done and provide details about the analyses in the comment box.**
   a) **Which quality assessment tool(s) were used in the review?** Check all that apply. (Risk of bias (RoB), Other (please name and describe),
   b) **Which methods of evidence synthesis were used in the review, besides a traditional narrative approach?** Check all that apply. (Meta-analysis, Qualitative synthesis, Other (please specify), No evidence synthesis methods, besides narrative)
   c) **Was a moderator or sub-group analysis conducted?** (Yes, No)

Population characteristics

6. **a) Which occupations are found in the studies actually included in the within-scope portion of the review? Check all that apply.** Refer to NOC 2016 v1.3

https://www23.statcan.gc.ca/imdb/p3VD.pl?Function=getVD&TVD=1267777 if needed. (Dentist/dentistry student; Pharmacist/pharmacy student; Nurse/nursing student (professional stream); Physician/medical resident/medical student; Other health occupations (in practice/students); describe briefly; Non-health occupations/students; describe briefly; Other, including any mixed groups; describe briefly; Unknown or missing)
b) Provide a summary statement for the occupation and registration status of all or most study participants across all studies.
c) What countries are the participants from in the studies included in the within-scope part of the review.

Intervention(s) characteristics

7. **Which (groups of) interventions were examined in the study?** *Note: These should be the (groups of) within-scope interventions the authors actually found and reported on.*

Outcome synthesis

8. **a) Identify and label all in-scope outcomes.** *Note: Use O1, O2, O3 etc. to label outcomes, e.g., O1 = post-training knowledge, O2 = Pre-post knowledge gain*.
b) Summarize the studies' intervention effects, relative to the control effects, for each outcome and place them under positive, negative or no effect sections below as appropriate. Outcome synthesis findings: (Positive Outcome (indicate n/a if not applicable); Negative Outcome (indicate n/a if not applicable); No effect (indicate n/a if not applicable))

Conclusions and Final Comment

9. **a) Summary of findings.** *Cut and paste synthesis/main findings of review. Include strength of evidence for each intervention category if reported*.
b) Additional "nuggets". *Enter here any "nuggets" of additional information you happen to notice and think are important to carry forward if we were to address the question "what works, how, in which conditions and for whom".*
10. **Enter any other notable information, e.g. about unusual features of the review that is not adequately captured in the other Data Extraction questions.**
11. **Is this the consensus – final - version of the DE form?** *Please select "no" until consensus has been achieved.* (Yes, No)

Appendix D: Quality assessment form, AMSTAR 2 (Level 5 form) – review 1

1. **Did the research questions and inclusion criteria for the review include the components of PICO**? *To score 'yes' appraisers should be confident that the 4 elements of PICO are described somewhere in the report.* (PICO component: Population, Intervention, Comparator group, Outcome).
   [Q1] **Optional (recommended)**. Time frame for follow-up.
   [Q1] Response *is auto populated based on the selections above.*
2. **Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?**
   (For partial Yes: *The authors state that they had a written protocol or guide that included ALL of the following:* review questions, a search strategy, inclusion/exclusion criteria, a risk of bias assessment.)
   [Q2] (For Yes: *As for partial yes, plus the protocol should be registered and should also have specified:* a meta-analysis/synthesis plan, if appropriate, and; a plan for investigating causes of heterogeneity; justification for any deviations from the protocol.)
   [Q2] Response is auto populated based on the selections above.
3. **Did the review authors explain their selection of the study designs (RCTs, or NRSIs, or both) for inclusion in the review?**
   (For Yes, the review should satisfy ONE of the following: Explanation for including only RCTs; OR Explanation for including NRSI; OR Explanation for including both RCTs and NRSI.)
   [Q3] Response is auto populated based on the selections above.
4. **Did the review authors use a comprehensive literature search strategy?**
   (For Partial Yes: The authors *state that they had a written protocol or guide that included ALL of the following:* searched at least 2 databases (relevant to research question); provided key word and/or search strategy; justified publication restrictions (e.g. language))
   [Q4] (For Yes: *As for partial yes, plus the protocol should be registered and should also have specified:* searched the reference lists/bibliographies of included studies; searched trial/study registries; included/consulted content experts in the field; where relevant, searched for grey literature; conducted search within 24 months of completion of the review.)
   [Q4] *Response is auto populated based on the selections above.*

5. **Did the review authors perform study selection in duplicate?** (*Choose either ONE of the following*: at least two reviewers independently agreed on selection of eligible studies and achieved consensus on which studies to include; OR two reviewers selected a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder selected by one reviewer.)
   [Q5] *Response is auto populated based on the selections above – YES if one selected.*
6. **Did the review authors perform data extraction in duplicate? (**Choose either ONE of the following: at least two reviewers achieved consensus on which data to extract from included studies; OR two reviewers extracted data from a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder extracted by one reviewer.
   [Q6] *Response is auto populated based on the selections above – YES if one selected.*
7. **Did the review authors provide a list of excluded studies and justify the exclusions?**
   (For partial yes: provided a list of all potentially relevant studies that were read in full text form but excluded from the review.)
   [Q7] (For Yes, must also have*:* justified the exclusion from the review of each potentially relevant study.)
   [Q7] *Response is auto populated based on the selections above.*
8. **Did the review authors describe the included studies in adequate detail?**
   (For Partial Yes (All of the following): described populations; described interventions; described comparators; described outcomes; described research designs.)
   [Q8] (For Yes, should also have ALL the following: described population in detail; described intervention and comparator in detail (including doses where relevant); described study's setting; timeframe for follow-up.)
   [Q8] *Response is auto populated based on the selections above.*
9. **RCTs- Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review**?
   (For Partial Yes, must have assessed RoB from: unconcealed allocation, and; lack of blinding of patients and assessors when assessing outcomes (unnecessary for objective outcomes such as all cause mortality).

[Q9 RCT] (For Yes, must also have assessed RoB from: allocation sequence that was not truly random, and; selection of the reported result from among multiple measurements or analyses of a specified outcome.)
[Q9 RCT] (Includes only NRSI.)
[Q9 RCT] *Response is auto populated based on the selections above.*

NRSI - Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?

(For Partial Yes, must have assessed RoB from: confounding, and; selection bias.)
[Q9 NRSI] (For Yes, must also have assessed RoB from: methods used to ascertain exposures and outcomes, and; selection of the reported result from among multiple measurements or analyses of a specified outcome.)
[Q9 NRSI] (Includes only RCT.)
[Q9 NRSI] *Response is auto populated based on the selections above.*

10. **Did the review authors report on the sources of funding for the studies included in the review?**
    (For Yes: Must have reported on the sources of funding for individual studies included in the review. Note: Reporting that the reviewers looked for this information but it was not reported by study authors also qualifies.)
    [Q10] *Response is auto populated based on the selections above.*
11. **RCTs - If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?**
    (For Yes, must also have assessed RoB from: The authors justified combining the data in a meta-analysis; AND they used an appropriate weighted technique to combine study results and adjusted for heterogeneity if present; AND investigated the causes of any heterogeneity.)
    [Q11 RCT] (No meta-analysis conducted.)
    [Q 11 RCT] *Response is auto populated based on the selections above.*

    NRSI- If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?

    (For Yes, must also have assessed RoB from: The authors justified combining the data in a meta-analysis; AND they used an appropriate weighted technique to combine study results, adjusting for heterogeneity if present; AND they statistically combined effect estimates from NRSI that

were adjusted for confounding, rather than combining raw data, or justified combining raw data when adjusted effect estimates were not available; AND they reported separate summary estimates for RCTs and NRSI separately when both were included in the review.)

[Q11 NRSI] (No meta-analysis conducted.)

[Q11 NRSI] *Response is auto populated based on the selections above.*

12. **If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?**
(For Yes: included only low risk of bias RCTs; OR, if the pooled estimate was based on RCTs and/or NRSI at variable RoB, the authors performed analyses to investigate possible impact of RoB on summary estimates of effect.)
[Q12] (No meta-analysis conducted.)
[Q12] *Response is auto populated based on the selections above.*
13. **Did the review authors account for RoB in individual studies when interpreting/discussing the results of the review?**
(For yes: included only low risk of bias RCTs; OR, if RCTs with moderate or high RoB, or NRSI were included the review provided a discussion of the likely impact of RoB on the results.)
[Q13] *Response is auto populated based on the selections above.*
14. **Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?**
(For Yes: There was no significant heterogeneity in the results; OR if heterogeneity was present the authors performed an investigation of sources of any heterogeneity in the results and discussed the impact of this on the results of the review.)
[Q14] *Response is auto populated based on the selections above.*
15. **If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?**
(For Yes: performed graphical or statistical tests for publication bias and discussed the likelihood and magnitude of impact of publication bias.)
[Q15] (No meta-analysis conducted.)
[Q15] *Response is auto populated based on the selections above.*
16. **Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?**

(For yes: The authors reported no competing interests OR; The authors described their funding sources and how they managed potential conflicts of interest.)

[Q16] *Response is auto populated based on the selections above.*

17. **Is this the consensus version?** Yes, No.
18. **Additional notes**

Appendix E: Reviews with e-learning as main focus, but no extractable results for SOL alone, excluded during full paper screening – review 1

Ahmadi, S.F., Baradaran, H.R., Ahmadi, E., 2015. Effectiveness of teaching evidence-based medicine to undergraduate medical students: a BEME systematic review. Medical Teacher 37, 21-30. doi:10.3109/0142159X.2014.971724

Al-Jewair, T.S., Qutub, A.F., Malkhassian, G., Dempster, L.J., 2010. A systematic review of computer-assisted learning in endodontics education. Journal of Dental Education 74, 601-611.

Alvarez, A.G., Dal Sasso, G.T.M., 2011. Virtual learning objects: contributions to the learning process in health and nursing. ACTA Paulista de Enfermagem 24, 707-711. doi:10.1590/s0103-21002011000500018

Ashokka, B., Dong, C., Law, L.S., Liaw, S.Y., Chen, F.G., Samarasekera, D.D., 2020. A BEME systematic review of teaching interventions to equip medical students and residents in early recognition and prompt escalation of acute clinical deteriorations: BEME Guide No. 62. Medical Teacher 42, 724-737. doi:10.1080/0142159X.2020.1763286

Bakkum, M.J., Tichelaar, J., Wellink, A., Richir, M.C., Van Agtmael, M.A., 2019. Digital learning to improve safe and effective prescribing: a systematic review. Clinical Pharmacology and Therapeutics 106, 1236-1245. doi:10.1002/cpt.1549

Bernard, R., Borokhovski, E., Schmid, R., Tamim, R., Abrami, P., 2014. A meta-analysis of blended learning and technology use in higher education: from the general to the applied. Journal of Computing in Higher Education 26, 87-122.

Botelho, M.G., Agrawal, K.R., Bornstein, M.M., 2019. An systematic review of e-learning outcomes in undergraduate dental radiology curricula-levels of learning and implications for researchers and curriculum planners. Dentomaxillofacial Radiology 48, 20180027. doi:10.1259/dmfr.20180027

Brusamento, S., Kyaw, B.M., Whiting, P., Li, L., Car, L.T., 2019. Digital health professions education in the field of pediatrics: systematic review and meta-analysis by the digital health education collaboration. Journal of Medical Internet Research 21, e14231. doi:10.2196/14231

Car, L.T., Myint Kyaw, B., Dunleavy, G., Smart, N.A., Semwal, M., Rotgans, J.I., Low-Beer, N., Campbell, J., 2019. Digital problem-based learning in health professions: systematic review and meta-analysis by the digital health education collaboration. Journal of Medical Internet Research 21, e12945. doi:10.2196/12945

Cook, D.A., Erwin, P.J., Triola, M.M., 2010. Computerized virtual patients in health

professions education: a systematic review and meta-analysis. Academic Medicine 85, 1589-1602. doi:10.1097/acm.0b013e3181edfe13

Cook, D.A., Hatala, R., Brydges, R., Zendejas, B., Szostek, J.H., Wang, A.T., Erwin, P.J., Hamstra, S.J., 2011. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA 306, 978-988. doi:10.1001/jama.2011.1234

Cook, D.A., Levinson, A.J., Garside, S., 2010. Time and learning efficiency in Internet-based learning: a systematic review and meta-analysis. Advances in Health Sciences Education 15, 755-770. doi:10.1007/s10459-010-9231-x

Cook, D.A., Levinson, A.J., Garside, S., Dupras, D.M., Erwin, P.J., Montori, V.M., 2010. Instructional design variations in internet-based learning for health professions education: a systematic review and meta-analysis. Academic Medicine 85, 909-922. doi:10.1097/ACM.0b013e3181d6c319

Du, S., Liu, Z., Liu, S., Yin, H., Xu, G., Zhang, H., Wang, A., 2013. Web-based distance learning for nurse education: a systematic review. International Nursing Review 60, 167-177. doi:10.1111/inr.12015

Frank, H.E., Becker-Haimes, E.M., Kendall, P.C., 2020. Therapist training in evidence-based interventions for mental health: a systematic review of training approaches and outcomes. Clinical Psychology 27, e12330.

Gao, Y., Gonzalez, V.A., Yiu, T.W., 2019. The effectiveness of traditional tools and computer-aided technologies for health and safety training in the construction sector: a systematic review. Computers & Education 138, 101-115. doi:10.1016/j.compedu.2019.05.003

Gentry, S.V., Gauthier, A., L'Estrade Ehrstrom, B., Wortley, D., Lilienthal, A., Tudor Car, L., Dauwels-Okutsu, S., Nikolaou, C.K., Zary, N., Campbell, J., Car, J., 2019. Serious gaming and gamification education in health professions: systematic review. Journal of Medical Internet Research 2019/03/29, e12994. doi:10.2196/12994

George, P.P., Papachristou, N., Belisario, J.M., Wang, W., Wark, P.A., Cotic, Z., Rasmussen, K., Sluiter, R., Riboli-Sasco, E., Car, L.T., Musulanov, E.M., Molina, J.A., Heng, B.H., Zhang, Y., Wheeler, E.L., Shorbaji, N.A., Majeed, A., Car, J., 2014. Online eLearning for undergraduates in health professions: a systematic review of the impact on knowledge, skills, attitudes and satisfaction. Journal of Global Health 4, 010406. doi:10.7189/jogh.04.010406

George, P.P., Zhabenko, O., Kyaw, B.M., Antoniou, P., Posadzki, P., Saxena, N., Semwal, M., Car, L.T., Zary, N., Lockwood, C., Car, J., 2019. Online digital education for postregistration training of medical doctors: systematic review by the digital health

education collaboration. Journal of Medical Internet Research 21, e13269. doi:10.2196/13269

Gordon, M., Patricio, M., Horne, L., Muston, A., Alston, S.R., Pammi, M., Thammasitboon, S., Park, S., Pawlikowska, T., Rees, E.L., Doyle, A.J., 2020. Developments in medical education in response to the COVID-19 pandemic: a rapid BEME systematic review: BEME Guide No. 63. Medical Teacher 42, 1202-1215. doi:10.1080/0142159X.2020.1807484

Harkanen, M., Voutilainen, A., Turunen, E., Vehvilainen-Julkunen, K., 2016. Systematic review and meta-analysis of educational interventions designed to improve medication administration skills and safety of registered nurses. Nurse Education Today 41, 36-43. doi:10.1016/j.nedt.2016.03.017

Ilic, D., Maloney, S., 2014. Methods of teaching medical trainees evidence-based medicine: a systematic review. Medical Education 48, 124-135. doi:10.1111/medu.12288

Kang, J., Seomun, G., 2018. Evaluating web-based nursing education's effects: a systematic review and meta-analysis. Western Journal of Nursing Research 40, 1677-1697. doi:10.1177/0193945917729160

Kononowicz, A.A., Woodham, L.A., Edelbring, S., Stathakarou, N., Davies, D., Saxena, N., Tudor Car, L., Carlstedt-Duke, J., Car, J., Zary, N., 2019. Virtual patient simulations in health professions education: systematic review and meta-analysis by the digital health education collaboration. Journal of Medical Internet Research 21, e14676. doi:10.2196/14676

Kyaw, B.M., Car, L.T., Galen, L.S.V., Van Agtmael, M.A., Costelloe, C.E., Ajuebor, O., Campbell, J., Car, J., 2019. Health professions digital education on antibiotic management: systematic review and meta-analysis by the digital health education collaboration. Journal of Medical Internet Research 21, e14984. doi:10.2196/14984

Kyaw, B.M., Saxena, N., Posadzki, P., Vseteckova, J., Nikolaou, C.K., George, P.P., Divakar, U., Masiello, I., Kononowicz, A.A., Zary, N., Car, L.T., 2019. Virtual reality for health professions education: systematic review and meta-analysis by the digital health education collaboration. Journal of Medical Internet Research 21, e12959. doi:10.2196/12959

Lahti, M., Hatonen, H., Valimaki, M., 2014. Impact of e-learning on nurses' and student nurses knowledge, skills, and satisfaction: a systematic review and meta-analysis. International Journal of Nursing Studies 51, 136-149. doi:10.1016/j.ijnurstu.2012.12.017

Lau, Y., Nyoe, R.S.S., Wong, S.N., Ab Hamid, Z.B., Leong, B.S.H., Lau, S.T., 2018.

Effectiveness of digital resuscitation training in improving knowledge and skills: a systematic review and meta-analysis of randomised controlled trials. Resuscitation 131, 14-23. doi:10.1016/j.resuscitation.2018.07.033

Liu, Q., Peng, W., Zhang, F., Hu, R., Li, Y., Yan, W., 2016. The effectiveness of blended learning in health professions: systematic review and meta-analysis. Journal of Medical Internet Research 18, e2. doi:10.2196/jmir.4807

Maheu-Cadotte, M.A., Cossette, S., Dube, V., Fontaine, G., Lavallee, A., Lavoie, P., Mailhot, T., Deschenes, M.F., 2020. Efficacy of serious games in healthcare professions education: a systematic review and meta-analysis. Simulation in Healthcare, [Online ahead of print]. doi:10.1097/sih.0000000000000512

Martinengo, L., Yeo, N.J.Y., Markandran, K.D., Olsson, M., Kyaw, B.M., Car, L.T., 2020. Digital health professions education on chronic wound management: a systematic review. International Journal of Nursing Studies 104, 103512. doi:10.1016/j.ijnurstu.2019.103512

McCall, M., Spencer, E., Owen, H., Roberts, N., Heneghan, C., 2018. Characteristics and efficacy of digital health education: an overview of systematic reviews. Health Education Journal 77, 497-514. doi:10.1177/0017896918762013

McCutcheon, K., Lohan, M., Traynor, M., Martin, D., 2015. A systematic review evaluating the impact of online or blended learning vs. face-to-face learning of clinical skills in undergraduate nurse education. Journal of Advanced Nursing 71, 255-270. doi:10.1111/jan.12509

McDonald, E.W., Boulton, J.L., Davis, J.L., 2018. E-learning and nursing assessment skills and knowledge: an integrative review. Nurse Education Today 66, 166-174. doi:10.1016/j.nedt.2018.03.011

Means, B., Toyama, Y., Murphy, R., Baki, M., 2013. The effectiveness of online and blended learning: a meta-analysis of the empirical literature. Teachers College Record 115, 1-47.

Militello, L.K., Gance-Cleveland, B., Aldrich, H., Kamal, R., 2014. A methodological quality synthesis of systematic reviews on computer-mediated continuing education for healthcare providers. Worldviews on Evidence-Based Nursing 11, 177-186. doi:10.1111/wvn.12041

Pastorino, R., Calabro, G.E., Lagerberg, T.B.M., Michelazzo, M.B., Boccia, S., 2018. Effectiveness of educational intervention types to improve genomic competency in non-geneticist medical doctors: a systematic review of the literature. Epidemiology Biostatistics and Public Health 15, e12657-12651. doi:10.2427/12657

Patel, R., Dennick, R., 2017. Simulation based teaching in interventional radiology training: is it effective? Clinical Radiology 72, 266.

Pei, L., Wu, H., 2019. Does online learning work better than offline learning in undergraduate medical education? A systematic review and meta-analysis. Medical Education Online 24, 1666538. doi:10.1080/10872981.2019.1666538

Phillips, E.A., Gordeev, V.S., Âgg, J., 2019. Effectiveness of occupational e-mental health interventions: a systematic review and meta-analysis of randomized controlled trials. Scandinavian Journal of Work, Environment and Health 45, 560-576. doi:10.5271/sjweh.3839

Piot, M.A., Dechartres, A., Attoe, C., Jollant, F., Lemogne, C., Layat Burn, C., Rethans, J.J., Michelet, D., Cross, S., Billon, G., Guerrier, G., Tesniere, A., Falissard, B., 2020. Simulation in psychiatry for medical doctors: a systematic review and meta-analysis. Medical Education 54, 696-708. doi:10.1111/medu.14166

Posadzki, P., Bala, M.M., Kyaw, B.M., Semwal, M., Divakar, U., Koperny, M., Sliwka, A., Car, J., 2019. Offline digital education for postregistration health professions: systematic review and meta-analysis by the digital health education collaboration. Journal of Medical Internet Research 2019/04/25, e12968. doi:10.2196/12968

Rao, G.K.L., Iskandar, Y.H., Mokhtar, N., 2020. Understanding the nuances of e-learning in orthodontic education. Education and Information Technologies 25, 307-328. doi:10.1007/s10639-019-09976-2

Rasmussen, K., Belisario, J.M., Wark, P.A., Molina, J.A., Loong, S.L., Cotic, Z., Papachristou, N., Riboli-Sasco, E., Tudor Car, L., Musulanov, E.M., Kunz, H., Zhang, Y., George, P.P., Heng, B.H., Wheeler, E.L., Al Shorbaji, N., Svab, I., Atun, R., Majeed, A., Car, J., 2014. Offline eLearning for undergraduates in health professions: a systematic review of the impact on knowledge, skills, attitudes and satisfaction. Journal of Global Health 4, 010405. doi:10.7189/jogh.04.010405

River, J., Currie, J., Crawford, T., Betihavas, V., Randall, S., 2016. A systematic review examining the effectiveness of blending technology with team-based learning. Nurse Education Today 45, 185-192. doi:10.1016/j.nedt.2016.08.012

Roh, K.H., Park, H.A., 2010. A meta-analysis on the effectiveness of computer-based education in nursing. Healthcare informatics research 16, 149-157. doi:10.4258/hir.2010.16.3.149

Rouleau, G., Gagnon, M.P., Cote, J., Payne-Gagnon, J., Hudson, E., Dubois, C.A., Bouix-Picasso, J., 2019. Effects of e-learning in a continuing education context on nursing care: systematic review of systematic qualitative, quantitative, and mixed-studies reviews. Journal of Medical Internet Research 21, e15118.

doi:10.2196/15118

Salter, S.M., Karia, A., Sanfilippo, F.M., Clifford, R.M., 2014. Effectiveness of E-learning in pharmacy education. American Journal of Pharmaceutical Education 78, 83. doi:10.5688/ajpe78483

Santos, G.N., Leite, A.F., Figueiredo, P.T., Pimentel, N.M., Flores-Mir, C., de Melo, N.S., Guerra, E.N., De Luca Canto, G., 2016. Effectiveness of e-learning in oral radiology education: a systematic review. Journal of Dental Education 80, 1126-1139.

Schmid, R.F., Bernard, R.M., Borokhovski, E., Tamim, R.M., Abrami, P.C., Surkes, M.A., Wade, C.A., Woods, J., 2014. The effects of technology use in postsecondary education: a meta-analysis of classroom applications. Computers & Education 72, 271-291.

Semwal, M., Whiting, P., Bajpai, R., Bajpai, S., Kyaw, B.M., Car, L.T., 2019. Digital education for health professions on smoking cessation management: systematic review by the digital health education collaboration. Journal of Medical Internet Research 21, e13000. doi:10.2196/13000

Sinclair, P.M., Kable, A., Levett-Jones, T., Booth, D., 2016. The effectiveness of Internet-based e-learning on clinician behaviour and patient outcomes: a systematic review. International Journal of Nursing Studies 57, 70-81. doi:10.1016/j.ijnurstu.2016.01.011

Tarpada, S.P., Hsueh, W.D., Gibber, M.J., 2017. Resident and student education in otolaryngology: a 10-year update on e-learning. Laryngoscope 127, E219-E224. doi:10.1002/lary.26320

Tarpada, S.P., Morris, M.T., Burton, D.A., 2016. E-learning in orthopedic surgery training: a systematic review. Journal of Orthopaedics 13, 425-430. doi:10.1016/j.jor.2016.09.004

Tomlinson, J., Shaw, T., Munro, A., Johnson, R., Madden, D.L., Phillips, R., McGregor, D., 2013. How does tele-learning compare with other forms of education delivery? A systematic review of tele-learning educational outcomes for health professionals. New South Wales Public Health Bulletin 24, 70-75.

Tudor Car, L., Soong, A., Kyaw, B.M., Chua, K.L., Low-Beer, N., Majeed, A., 2019. Health professions digital education on clinical practice guidelines: a systematic review by Digital Health Education collaboration. BMC Medicine 17, 139. doi:10.1186/s12916-019-1370-1

Vallee, A., Blacher, J., Cariou, A., Sorbets, E., 2020. Blended learning compared to

traditional learning in medical education: systematic review and meta-analysis. Journal of Medical Internet Research 22, e16504. doi:10.2196/16504

Veneri, D., 2011. The role and effectiveness of computer-assisted learning in physical therapy education: a systematic review. Physiotherapy Theory and Practice 27, 287-298. doi:10.3109/09593985.2010.493192

Viljoen, C.A., Scott Millar, R., Engel, M.E., Shelton, M., Burch, V., 2019. Is computer-assisted instruction more effective than other educational methods in achieving ECG competence amongst medical students and residents? A systematic review and meta-analysis. BMJ Open 9, e028800. doi:10.1136/bmjopen-2018-028800

Voutilainen, A., Saaranen, T., Sormunen, M., 2017. Conventional vs. e-learning in nursing education: a systematic review and meta-analysis. Nurse Education Today 50, 97-103. doi:10.1016/j.nedt.2016.12.020

Wahabi, H.A., Esmaeil, S.A., Bahkali, K.H., Titi, M.A., Amer, Y.S., Ahmed, A., Jamal, A., Zakaria, N., Siddiqui, A.R., Semwal, M., Car, L.T., Posadzki, P., Car, J., 2019. Medical doctors' offline computer-assisted digital education: systematic review by the digital health education collaboration. Journal of Medical Internet Research 21, e12998. doi:10.2196/12998

Webb, L., Clough, J., O'Reilly, D., Wilmott, D., Witham, G., 2017. The utility and impact of information communication technology (ICT) for pre-registration nurse education: a narrative synthesis systematic review. Nurse Education Today 48, 160-171. doi:10.1016/j.nedt.2016.10.007

Wilson, A.B., Brown, K.M., Misch, J., Miller, C.H., Klein, B.A., Taylor, M.A., Goodwin, M., Boyle, E.K., Hoppe, C., Lazarus, M.D., 2019. Breaking with tradition: a scoping meta-analysis analyzing the effects of student-centered learning and computer-aided instruction on student performance in anatomy. Anatomical Sciences Education 12, 61-73. doi:10.1002/ase.1789

Woldeab, D., Yawson, R.M., Osafo, E., 2020. A systematic meta-analytic review of thinking beyond the comparison of online versus traditional learning. E-Journal of Business Education & Scholarship of Teaching 14, 1-24.

Xu, X., Posadzki, P.P., Lee, G.E., Car, J., Smith, H.E., 2019. Digital education for health professions in the field of dermatology: a systematic review by digital health education collaboration. Acta Dermato-Venereologica 99, 133-138. doi:10.2340/00015555-3068

Zafar, S., Safdar, S., Zafar, A.N., 2014. Evaluation of use of e-Learning in undergraduate radiology education: a review. European Journal of Radiology 83, 2277-2287.

Zhao, J., Xu, X., Jiang, H., Ding, Y., 2020. The effectiveness of virtual reality-based technology on anatomy teaching: a meta-analysis of randomized controlled studies. BMC Medical Education 20, 127. doi:10.1186/s12909-020-1994-z

Appendix F: List of systematic reviews included in review I

Chipps, J., Brysiewicz, P., Mars, M., 2012. A systematic review of the effectiveness of videoconference-based tele-education for medical and nursing education. Worldviews on Evidence-Based Nursing 9, 78-87. doi:10.1111/j.1741-6787.2012.00241.x

Gegenfurtner, A., Ebner, C., 2019. Webinars in higher education and professional training: a meta-analysis and systematic review of randomized controlled trials. Educational Research Review 28, 100293. doi:10.1016/j.edurev.2019.100293

He, L., Yang, N., Xu, L., Ping, F., Li, W., Sun, Q., Li, Y., Zhu, H., Zhang, H., 2020. Synchronous distance education vs traditional education for health science students: a systematic review and meta-analysis. Medical Education 55, 293-308. doi:10.1111/medu.14364

Appendix G: Search strategy for MEDLINE – review 2

16 December 2020
N=2022

Database: Ovid MEDLINE: Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE® Daily and Ovid MEDLINE® <1946-Present>
Search Strategy:
---
1 exp Education, Distance/ (4381)
2 exp Computer-Assisted Instruction/ (11977)
3 (distance adj2 learning).ti,ab. (1431)
4 (distance adj2 education).ti,ab. (939)
5 (distance adj2 instruct*).ti,ab. (28)
6 (distance adj2 teach*).ti,ab. (75)
7 (distance adj2 training).ti,ab. (366)
8 e-learning.ti,ab. (2594)
9 elearning.ti,ab. (249)
10 (e-training or e-trainer$).ti,ab. (71)
11 (etraining or etrainer$).ti,ab. (2)
12 "online classroom*".ti,ab. (45)
13 (online adj2 course$).ti,ab. (1371)
14 (online adj2 education).ti,ab. (1242)
15 (online adj2 learning).ti,ab. (1970)
16 (online adj2 lecture$).ti,ab. (189)
17 (online adj2 instruct*).ti,ab. (1682)
18 (online adj2 teach*).ti,ab. (469)
19 (online adj2 training).ti,ab. (1060)
20 (remote adj2 instruct*).ti,ab. (45)
21 (remote adj2 learning).ti,ab. (127)
22 (remote adj2 teach*).ti,ab. (74)
23 (remote adj2 training).ti,ab. (139)
24 "virtual classroom*".ti,ab. (132)
25 (virtual* adj2 learning).ti,ab. (662)
26 (virtual* adj2 instruct*).ti,ab. (56)
27 (virtual* adj2 live).ti,ab. (58)
28 (virtual adj2 teach*).ti,ab. (102)
29 (virtual* adj2 train*).ti,ab. (1029)
30 "web-based training".ti,ab. (287)
31 (synchronous adj2 learning).ti,ab. (46)
32 or/1-31 (25592)
33 exp child/ (1934626)
34 "K-12".ti,ab. (8284)
35 ("elementary school$" or "middle school$" or "junior high" or "high school$" or "secondary school$").ti,ab. (57403)

36 33 or 34 or 35 (1982601)
37 32 not 36 (24232)
38 accident$.ti,ab. (115074)
39 exp Accidents/ or exp Accidents, Occupational/ (193399)
40 exp Occupational Health/ (34120)
41 best practices.ti,ab. (14026)
42 behavio?r.ti,ab. (855323)
43 claim$.ti,ab. (96364)
44 Evaluation study/ (255504)
45 Health Behavior/ (51224)
46 injur*.ti,ab. (822189)
47 knowledge.ti,ab. (727515)
48 exp Occupational Diseases/ (133339)
49 exp Occupational Exposure/ (64429)
50 (occupation* adj2 exposure).ti,ab. (20978)
51 practice.ti,ab. (725759)
52 safety.ti,ab. (527951)
53 Safety Management/ (20403)
54 (work* adj2 engag*).ti,ab. (3744)
55 (work* adj2 exposure).ti,ab. (6025)
56 (work adj2 participat*).ti,ab. (1147)
57 (work* adj3 practice*).ti,ab. (12331)
58 exp Workers' Compensation/ (7548)
59 "Wounds and Injuries"/ (77851)
60 skill$.ti,ab. (205933)
61 survey$.ti,ab. (668419)
62 test$.ti,ab. (3292191)
63 score$.ti,ab. (969716)
64 grade$.ti,ab. (435221)
65 learning.ti,ab. (311637)
66 or/38-65 (8067300)
67 37 and 66 (17623)
68 67 (17623)
69 limit 68 to english language (16954)
70 limit 69 to yr="2020" (2022)

Appendix H: Screening Forms (Levels 1, 2) – review 2

Screening Form – Level 1

1. **Does this report involve a "virtual classroom" education/training of an occupational/professional/career-preparing nature?** (Yes, No, Can't Tell)
2. **Does this report involve a comparison of "virtual classroom" with another e-learning or face-to-face classroom training?** (Yes, No, Can't Tell)
3. **Are study participants adults in a developed country?** (Yes (include); No (exclude); Can't tell (include))
4. **Is one or more eligible outcomes reported upon?** (Yes, No, Can't Tell)
5. **Are there other reasons to flag this record for possible use?** (Conceptual contribution, Review article, Relevant qualitative study, Protocol for future eligible study, Other)

Screening Form – Level 2

1. **Does this report involve a primary study of a "virtual classroom" education/training of an occupational/professional/career-preparing nature?** (Yes, No)
2. **Are study participants adults in a developed country?** (Yes, No)
3. **Is one or more eligible outcomes reported upon?** (Yes, No)
4. **Does this report involve a comparison of "virtual classroom" with another e-learning or face-to-face classroom training?** (Yes, virtual classroom vs face-to-face classroom; Yes, virtual classroom vs other e-learning; No comparison of "virtual classroom" with another e-learning or face-to-face classroom training)
5. **Are there other reasons to flag this record for possible use?** (Conceptual contribution, Review article, Relevant qualitative study, Protocol for future eligible study, OHS study, Other)

Appendix I: Quality assessment form (Level 3) – review 2

Check on Relevancy

1. **Should this article continue to quality appraisal?** (Yes, the article is relevant and should proceed to quality appraisal; No, the article is NOT relevant and should NOT proceed to quality appraisal)
2. **Why do you believe this study does not meet the inclusion criteria?** *(If you answered 'No' to Q1, please provide the specific reason(s) for why you do not feel this study meets the inclusion criteria. When answering, please select all relevant responses and provide an explanation in the comment field(s))* (The publication type is not eligible, There is no virtual classroom intervention, There is no comparison of virtual classroom with face-to-face classroom or other e-learning, The outcome(s) described are not related to Kirkpatrick level 2, 3 or 4, The population does not include adults from a developed country, Other (please elaborate))

Design and objectives

3. **What is the study design?** (Randomized controlled trial, Nonrandomized controlled trial, Before-and-after study/pre-post (no control), Cohort study, Cross-sectional study, Other (please specify in the text box))
4. **Is the research question/objective clearly stated?** (Yes = 2, No = 0)

Recruitment

5. **Were sampling and recruitment methods (including inclusion/exclusion criteria) clearly described and similar for all participants?** (Yes = 2, No = 0)
6. **Was recruitment (or participant) rate reported and adequate?** (Yes, rates were equal to and greater than 70% = 2; No, participation rate was less than 55% = 0; Participation was not reported = 0; Not applicable => item excluded from scoring)
7. **Were there important differences between those who participated and did not participate in the study with respect to key characteristics i.e., exposure/intervention(s) and/or outcome(s)?** *Note that the response should be "Not Applicable" if there was a 100% participation rate for the study or if Q6 was answered as "Not applicable"*. (Yes, described and no major differences = 2; Yes, described and there were major differences (explain in

comment box) = 0; Not described or assessed = 0; Not applicable => item excluded from scoring)

Intervention characteristics

8. **Was an intervention allocation method performed adequately?** *To score yes, the method of randomization AND allocation concealment have to be adequate. Note that the response should be "Note applicable" if this was an observational study (i.e., the investigators had no control over the allocation of the exposure/intervention) or there was no control group*. (Yes, adequately = 2; No, not adequately = 0; No, not described = 0; Not applicable => item excluded from scoring)
9. **Was the intervention process adequately described to allow for replication?** (Yes, adequately = 2; No, not adequately = 0; No, not described = 0)
10. **Was there any potential for contamination?** *Note that the response should be "Not applicable" if there was no control group.* (Yes and not accounted for = 0; No or at least accounted for = 2; Not described = 0 Not applicable => item excluded from scoring)
11. **Was there any potential for co-intervention?** (Yes and not accounted for = 0; No or at least accounted for = 2; Not described = 0)
12. **Was compliance with the intervention described and adequate?** *If there are differences in compliance across different groups, please clarify in the text box*. (Yes, adequate or at least accounted for = 2; No, not adequate and not accounted for = 0; Not described =0)
13. **Was the length-of-follow up (from baseline) appropriate for the intervention and outcomes?** *Note that the response should be "Not applicable" if the study was a cross-sectional study where there was no follow-up.* (yes = 2; No = 0; Not described = 0; Not applicable => item excluded from scoring)

Attrition bias

14. **Was the loss to follow up (attrition) less than 35%?** *Note that the response should be "Not applicable" if the analysis was based solely on administrative or existing data sources (i.e., did not require active recruitment of participants) or the study was not longitudinal in nature (i.e., did not actively follow up a sample)*. (Yes, less than 35% = 2; No, more than or equal to 35% = 0; Not described = 0; Not applicable => item excluded from scoring)

15. **Were there important differences between those who completed the study and those who withdrew with respect to key characteristics i.e., intervention(s) and outcome(s)?** *Note that the response should be "Not applicable" if there was a 100% retention rate for the study or if Q13 was answered as "Not applicable".* (Yes, described and no major differences between groups = 2; Yes, described and there were major differences between groups (explain in comment box) = 0; Not described or assessed = 0; Not applicable => item excluded from scoring

Outcome measurement bias

16. **Were the outcomes described at baseline and follow-up?** *Note that the response should be "Not applicable" if the study was a cross-sectional study where there was no follow-up*. (Yes, described at baseline and follow-up = 2; No, only described at baseline = 1; No, only described at follow-up = 1; Not described or assessed = 0; Not applicable => item excluded from scoring)
17. **Were the instruments/methods used to assess the outcome(s) valid, reliable, and not prone to important sources of measurement bias?** *For studies that use a mix of measurement methods, where some may be considered valid and reliable and others not, please select "Partial" as a response. Note: For studies that look at various outcomes, please ensure that you answer this question only for outcomes that are relevant for our research question*. (Yes = 2; No = 0; Unable to determine = 0; Not adequately described = 0; Partial = 0)
18. **Was data collection with respect to outcome carried out equivalently for all participants?** *Note: For studies that look at various outcomes, please ensure that you answer this question only for outcomes that are relevant for our research question*. (Yes = 2; No = 0)

Confounding and analysis

19. **Were baseline characteristics described?** (Yes, described = 2; No, not described = 0)
20. **Were all participants' outcomes analyzed by the groups to which they were originally allocated (intention-to-treat analysis)?** *Note that the response should be "Not applicable" if this was an observational study (i.e., the investigators had no control over the allocation of the exposure/intervention)*. (Yes = 2; No = 0; Not described = 0; Not applicable - observational study => item excluded from scoring)

21. **Were important covariates, confounders, or baseline differences (if necessary) accounted for in the study design and/or analysis?** (Yes or no differences at baseline = 2; Partially = 1; No = 0; Not applicable => item excluded from scoring)
22. **Was there a direct between group comparison?** (Yes = 2; No = 0)
23. **Is there anything you'd like to note about the study? This can be a fatal flaw, something to note, or your overall impressions?**

Item weights (purpose described in Methods section). Items with weight = 1: 4, 9, 14, 19, 22. Items with weight = 2: 6, 10, 11, 12, 17. Items with weight = 3: 5, 7, 8, 13, 15, 16, 18, 20, 21.

Appendix J: Data extraction form (Level 4) – review 2

1. **Should this article be excluded from DE because it does not meet our inclusion criteria for the population, intervention, comparison and outcomes?** (Yes (please indicate why), No)

Study design and Setting

2. **State the research question/objective(s).**
3. **Write the last name of the first author and the year of publication.** (First author's last name, Year of publication)
4. **State the country/countries where the study was completed.**
5. **Over what time period was the study conducted?**
6. **Describe the source population from which the participants were recruited?** *Please include details such as job titles, year of studies for students, etc*.
7. **What is the largest analytical sample size across the outcomes reported below?**
8. **Describe the type of setting the study was conducted in, if applicable.** *Please copy and paste these details straight from the paper*.
9. **Please clearly list the inclusion and exclusion criteria described in the study.** *Please copy-paste all details for both inclusion and exclusion criteria separately*.
10. **What was the study design?** (a) Randomized controlled trial, b) Non-randomized study (pre-post), c) Before-and-after study / pre-post (no control), d) Cohort study, e) Cross sectional study, f) Other (please specify))

Intervention characteristics

Intervention group

11. **Describe the field of intervention for the intervention group**.
12. **Describe the learning topics/objectives of the intervention for the intervention group**.
13. **Describe the intervention components for the intervention group**.
14. **Describe the number of sessions for the intervention among the intervention group**.
15. **Describe the total learning time for the intervention among the intervention group**.
16. **Describe the delivery period for the intervention group**.

Control group

17. **Describe the field of intervention for the control group**.
18. **Describe the learning topics/objectives of the intervention for the control group**.
19. **Describe the intervention components for the control group**.
20. **Describe the number of sessions for the intervention among the control group**.
21. **Describe the total learning time for the intervention among the control group**.
22. **Describe the delivery period for the control group**.

Outcomes

23. **Provide a list of outcome variables used to evaluate intervention effectiveness that are relevant to our review project, as well as every time point at which each outcome variable was examined**.
24. **Describe, for each outcome of interest, the outcome measurement(s)**.
25. **Describe, for each outcome of interest, the observed intervention effects**.
26. **List all covariates/confounders that were controlled for in the final analysis and describe whether they were the same for each outcome. If there were differences across outcomes, please provide details**.
27. **Remark on the findings or enter information that is unique about the study that may not be adequately captured in the other DE questions**.

Appendix K: List of included studies – review 2

Carroll, T., Mooney, C., Horowitz, R., 2021. Re-ACT: Remote Advanced Communication Training in a time of crisis. Journal of Pain and Symptom Management 61, 364-368. doi:10.1016/j.jpainsymman.2020.08.013

Dahlstrom-Hakki, I., Alstad, Z., Banerjee, M., 2020. Comparing synchronous and asynchronous online discussions for students with disabilities: the impact of social presence. Computers & Education 150, 103842. doi:10.1016/j.compedu.2020.103842

Dela Cruz, A.M., Alick, S., Das, R., Brenner, A., 2020. Same material, different formats: comparing in-person and distance learning in undergraduate medical education. Academic Psychiatry 14, 659-663. doi:10.1007/s40596-020-01333-7

Elliott, G.M., Henninger, J., 2020. Online teaching and self-efficacy to work with suicidal clients. Counselor Education & Supervision 59, 283-296. doi:10.1002/ceas.12189

Guo, S., 2020. Synchronous versus asynchronous online teaching of physics during the COVID-19 pandemic. Physics Education 55, 1-9. doi:10.1088/1361-6552/aba1c5

Lee, S.J., Park, J., Lee, Y.J., Lee, S., Kim, W.H., Yoon, H.B., 2020. The feasibility and satisfaction of an online global health education course at a single medical school: a retrospective study. Korean Journal of Medical Education 32, 307-315. doi:10.3946/kjme.2020.178

Morice, A., Jablon, E., Delevaque, C., Khonsari, R.H., Picard, A., Kadlub, N., 2020. Virtual versus traditional classroom on facial traumatology learning: evaluation of medical student's knowledge acquisition and satisfaction. Journal of Stomatology, Oral and Maxillofacial Surgery 121, 642-645. doi:10.1016/j.jormas.2020.03.001

Mullen, C.A., 2020. Does modality matter? A comparison of aspiring leaders' learning online and face-to-face. Journal of Further and Higher Education 44, 670-688. doi:10.1080/0309877X.2019.1576859

Norton, S., 2020. Australian primary mathematics teacher preparation: on-campus or online? Who? Why? So What? Mathematics Teacher Education & Development 22, 91-114.

Root, W.B., Rehfeldt, R.A. 2021. Towards a modern-day teaching machine: the synthesis of programmed instruction and online education. The Psychological Record 71(1): 85-94.

Supplementing Root & Rehfeldt 2021 (additional information on same research studies):

Root, W.B., 2019. The synthesis of programmed instruction and online education: towards a modern-day teaching machine [PhD thesis]. Rehabilitation Department, Southern Illinois University, Carbondale, IL.

Appendix L: Quality assessment results – detailed – review 2

Q #	QA Question (full questions in Appendix I)	Wt	Ref#1	Ref# 2	Ref#3	Ref#4	Ref#5	Ref#6	Ref# 7	Ref#8	Ref#9	Ref#10
3	... study design?	3	-	-	-	-	-	-	-	-	-	-
4	… research question clearly stated?	1	Y=2	Y=2	N=0	Y=2	Y=2	Y=2	Y=1	Y=1	Y=2	Y=2
5	… sampling and recruitment methods clearly described and similar for all participants?	3	N=0	Y=2	N=0	Y=2	Y=2	Y=2	N=0	Y=2	Y=2	N=0
6	…recruitment (or participation) rate reported and adequate?	2	N=0	nr=0	Y=2	nr=0	N/A	N/A	P=1	Y=2	N/A	nr=0
7	… important differences between those who participated and did not participate … with respect to key characteristics …?	3	nr=0	nr=0	nr=0	nr=0	N/A	N/A	nr=0	nr=0	N/A	N/A
8	… intervention allocation method performed adequately?	3	N/A	Y=2	N/A	Y=2	N/A	N/A	N/A	N/A	N/A	Y=2
9	… intervention process adequately described to allow for replication?	1	Y=2	nr=0	Y=2	Y=2	Y=2	Y=2	Y=2	Y=2	Y=2	Y=2
10	… any potential for contamination?	2	nr=0	nr=0	N/AC=2	nr=0	N/AC=2	N/AC=2	N/AC=2	N/AC=2	Y=0	N/AC=2
11	… any potential for co-intervention?	2	Y=0	N/AC=2	Y=0	N/AC=2	nr=0	N/AC=2	Y=0	N/AC=2	N/AC=2	N/AC=2
12	… compliance with the intervention described and adequate?	2	nr=0	nr=0	Y=2	nr=0	nr=0	Y=2	nr=0	nr=0	N=0	nr=0
13	… length-of-followup … appropriate for the intervention and outcomes?	3	nr=0	nr=0	Y=2	Y=2	Y=2	Y=2	Y=2	Y=2	Y=2	Y=2
14	… loss to follow up … less than 35%?	1	N=0	nr=0	nr=0	N=0	<35%=2	nr=0	<35%=2	<35%=2	<35%=2	nr=0
15	… important differences between those who completed the study and those who withdrew …?	3	nr=0	nr=0	nr=0	nr=0	nr=0	nr=0	nr=0	N/A	nr=0	nr=0
16	… outcomes described at baseline and follow-up?	3	FU=1	BFU=2	FU=1	BFU=2	FU=1	FU=1	FU=1	FU=1	FU=1	FU=1
17	… methods used to assess the outcome(s) valid, reliable…?	2	nr=0	Y=2	nr=0	Y=2	N=0	Y=2	Y=2	nr=0	nr=0	nr=0

Q #	QA Question (full questions in Appendix I)	Wt	Ref#1	Ref# 2	Ref#3	Ref#4	Ref#5	Ref#6	Ref# 7	Ref#8	Ref#9	Ref#10
18	… [outcome] data collection … carried out equivalently for all participants?	3	N=0	Y=2	N=0	Y=2	N=0	N=0	N=0	Y=2	Y=2	Y=2
19	… baseline characteristics described?	1	N=0	Y=2	N=0	Y=2	N=0	N=0	N=0	Y=2	Y=2	Y=2
20	… outcomes analyzed by the groups to which they were originally allocated …?	3	N/A	Y=2	N/A	Y=2	N/A	N/A	N/A	N/A	N/A	Y=2
21	… important covariates, confounders, or baseline differences … accounted for in the study design and/or analysis?	3	N=0	Y=2	N=0	Y=2	N=0	N=0	N=0	N=0	N=0	Y=2
22	… direct between-group comparison?	1	N=0	Y=2	N=0	Y=2	Y=2	Y=2	Y=2	Y=2	Y=2	Y=2
	Total raw score		**5**	**22**	**11**	**26**	**15**	**19**	**16**	**21**	**19**	**23**
	Total weighted score		**7**	**50**	**23**	**58**	**27**	**37**	**27**	**41**	**35**	**49**
	Denominator		**72**	**84**	**72**	**84**	**62**	**62**	**72**	**66**	**62**	**78**
	Final score		**9.7**	**59.5**	**31.9**	**69**	**43.5**	**59.7**	**37.5**	**62.1**	**56.5**	**62.8**

Total raw score: sum of raw scores (displayed in table) for all items applicable to the reference; a non-applicable item is indicated by N/A and is excluded from the scoring. *Total weighted score*: sum of weighted scores for all applicable items; a weighted score for an item is the raw score multiplied by the weight in column 3. *Denominator* is computed from sum of maximum possible weighted scores for all applicable items.
Abbreviations: BFU=Baseline and follow up; FU=Follow up only; N=No; N/A= Not applicable; N/AC= No or accounted for; nr= not reported, not described or unable to determine; P=Partial; Wt=Importance weight based on team consensus; Y= yes. Key to references: Ref#1= Carroll et al., 2021; Ref#2= Dahlstrom-Hakki et al.,, 2020; Ref#3= Dela Cruz et al., 2020; Ref#4= Elliott & Henninger, 2020; Ref#5= Guo, 2020; Ref#6= Lee et al., 2020; Ref#7= Morice et al., 2020; Ref#8= Mullen, 2020; Ref#9= Norton, 2020; Ref#10= Root & Rehfeldt, 2021.

Appendix M: Detailed results on the relative effectiveness of SOL versus F2F learning on post-training knowledge – review 2

Study 1st author, year published	Study design n total	Outcome measurement	Effects of intervention, reported			Author's conclusion about effects (p-value for between-treatment test)
			SOL treatment	F2F treatment	Between treatment effects/tests	
Lee, 2020	Nonconcurrent cohort 299	Final exam score (%), mean (sd) Note: Percentage of exam score from multiple choice to percentage from essay: 2019: 80:20 2020: 70:30	84.18 (19.63)	78.04 (18.33)	Independent t-test statistic of SOL vs F2F difference: -2.82 (p<0.005)	SOL > F2F (p<0.005)
Mullen, 2020	Cohort 14	Cognitive presence, indicator of adult learning, measured by # of pages in 4 submitted assignments (# as proxy for quality): O1: RPP x 4 O2: CAPP O3: APAP O4: LFSP Total (calculated by reviewers)	O1: 228 O2: 152 O3: 142 O4: 5 Total: 524	O1: 266 O2: 132 O3: 141 O4: 7 Total: 546	F2F generated 18 more pages, a difference small enough that the author considered the 2 groups "equal"	SOL = F2F (nr)

Study 1st author, year published	Study design n total	Outcome measurement	Effects of intervention, reported			Author's conclusion about effects (p-value for between-treatment test)
			SOL treatment	F2F treatment	Between treatment effects/tests	
Norton, 2020	Cohort 342	Content knowledge of math and math pedagogy exam scores out of 100, mean (sd) O1: Exam 1 (math for lower primary grades) O2: Exam 2 (math for upper primary grades)	O1: 62.88 (13.89) O2: 59.38 (17.56)	O1: 59.72 (13.55) O2: 53.52 (16.88)	ANOVA tests of SOL vs F2F differences: O1: p = 0.128 O2: p = 0.035	F2F not more effective than SOL (O2 differences attributed to pre-existing group differences, including prior math training)
Root, 2021; E1	Single case, alternating treatment 24	% correct in weekly quizzes, mean of within-subject means for 4 or 5 treatment phases (sd)	*Reviewers:* 83.7 (7.2)	*Reviewers:* 77.2 (9.3)	Within-subject average SOL-F2F difference (sd): *Reviewers:* 6.5 (7.5) Paired t-test: *Reviewers:* t(23)=4.24 p = 0.0003	*Reviewers:*[1] SOL > F2F P = 0.0003

#, number; E1, experiment 1; nr, not reported; O1, O2 = outcomes 1 and 2; sd, standard deviation

[1] Reviewers conducted own statistical calculations, using data from Table 1 Root (2019), with permission of Dr. Root.

Appendix N: Detailed results found on the relative effectiveness of SOL versus other e-learning on knowledge and self-efficacy outcomes – review 2

Study 1st author, year published	Outcome(s)	Outcome measurement	Effects of intervention, reported			Author's conclusion about effects (p-value for between-treatment test)
			SOL treatment	Comparison treatment	Between treatment	
Dahlstrom-Hakki, 2020	Knowledge-pre-post	Test score, mean (sd)	SOL: Pre: 2.43 (1.52) Post: 2.46 (1.57)	AEL: Pre: 2.40 (1.52) Post: 2.68 (1.61)	Pretest x SOL/AEL interaction variable in GLMEM model ($p < 0.1$)	SOL= AEL ($p < 0.1$)
Elliott, 2020	Self-efficacy-pre-post	Counselor Suicide Assessment Efficacy Survey: 25-items, responses from 1 = not confident to 5 = highly confident, overall scores range 25 to 125.	Within-group paired-samples t tests: I1: $t(19) = –5.95$, $p < .001$ I2: $t(32) = –7.56$, $p < .001$		ANCOVA among I1, I2 and no-SOL control group: $F(2,84)=0.03$, $p=0.97$, partial $\eta^2=0.001$	I1 (role-playing) = I2 (Q & A) ($p = 0.97$)

Study 1st author, year published	Outcome(s)	Outcome measurement	Effects of intervention, reported			Author's conclusion about effects (p-value for between-treatment test)
			SOL treatment	Comparison treatment	Between treatment	
Root, 2021; E2[1]	Knowledge-post	% correct, mean of within-subject mean for multiple treatment phases (sd)	Reviewers: 89.6 (8.4)	Reviewers: 91.1 (3.9)	Reviewers: Within-subject average difference (sd) -0.15 (9.7) Paired t-test: t(7) = -0.43, p =0.68	Reviewers: SOL = AEL (p = 0.68)
Root, 2021; E3[2]	Knowledge-post	% correct, mean of within-subject mean for multiple treatment phases (sd)			Reviewers: Within-subject average I1-I2 difference (sd): -1.65 (8.60) Paired t-test t(9) = -0.61 p = 0.56	Reviewers: I1 (SOL lecture + AEL discussion) = I2 (AEL modules + SOL chat) (p = 0.56)

Abbreviations: AEL, asynchronous online learning; E2 and E3, experiments 2 and 3; I1 and I2, interventions 1 and 2; sd, standard deviation; SOL, synchronous online learning

[1] Reviewers' calculations based on raw data in Table 4 of Root (2019), with permission of Dr. Root.

[2] Reviewers' calculations based on values estimated from Fig. 3 of Root & Rehfeldt (2021), with permission of Dr. Root.

Appendix O: OHS-related articles excluded from review 2 at the screening stage

1. Akanmu, A.A., Olayiwola, J., Ogunseiju, O., McFeeters, D., 2020. Cyber-physical postural training system for construction workers. Automation in Construction 117, 103272. DOI: 10.1016/j.autcon.2020.103272
2. Aragão, R., Pereira-Guizzo, C., Figueiredo, P.S., 2020. Related information Impacts of an e-learning system on the occurrence of work accidents in a chemical industry company. International Journal of Knowledge Management Studies 11 (4), 325 – 343. DOI: 10.1504/IJKMS.2020.110667
3. Dhalmahapatra, K., Das, S., and Maiti, J. 2020. On accident causation models, safety training and virtual reality. International Journal of Occupational Safety and Ergonomics, DOI: 10.1080/10803548.2020.1766290.
4. Eiris, R., Gheisari, M., Esmaeili, B., 2020. Desktop-based safety training using 360-degree panorama and static virtual reality techniques: A comparative experimental study. Automation in Construction 109, 102969. DOI: 10.1016/j.autcon.2019.102969.
5. Eiris, R., Jain, A., Gheisari, M., Wehle, A., 2020. Safety immersive storytelling using narrated 360-degree panoramas: A fall hazard training within the electrical trade context. Safety Science 127, 104703. DOI: 10.1016/j.ssci.2020.104703.
6. Joshi, S., Hamilton, M., Warren, R., Faucett, D., Tian, W., Wang, Y., Ma, J., 2021. Implementing Virtual Reality technology for safety training in the precast/ prestressed concrete industry. Applied Ergonomics 90, 103286. DOI: 10.1016/j.apergo.2020.103286
7. Lanzotti, A., Vanacore, A., Tarallo, A., Nathan-Roberts, D., Coccorese, D., Minopoli, V., Carbone, F., d'Angelo, R., Grasso, C., Di Gironimo, G., and Papa, S., 2020. Interactive tools for safety 4.0: virtual ergonomics and serious games in real working contexts. Ergonomics 63(3), 324 – 333. DOI: 10.1080/00140139.2019.1683603.
8. Nykänen, M., Puro, V. Tiikkaja, M., Kannisto, H., Lantto, E., Simpura, F., Uusitalo, J., Lukander K., Räsänen, T. Heikkilä, T., Teperi, A.-M., 2020. Implementing and evaluating novel safety training methods for construction sector workers: results of a randomized controlled trial. Journal of Safety Research 75, 205-221, DOI: 10.1016/j.jsr.2020.09.015.
9. Pedram, S., Palmisano, S., Skarbez, R., Perez, P., Farrelly, M., 2020. Investigating the process of mine rescuers' safety training with immersive virtual reality: A structural equation modelling approach. Computers & Education 153, 103891. DOI: 10.1016/j.compedu.2020.103891

10. Saghafian, M., Laumann, K., Akhtar, R.S., Skogstad, M.R., 2020. The evaluation of virtual reality fire extinguisher training. Frontiers in Psychology 11, 593466. DOI: 10.3389/fpsyg.2020.593466

11. Suppan, L., Abbas, M., Stuby, L., Cottet, P., Larribau, R., Golay, E., Iten, A., Harbarth, S., Gartner, B., Suppan, M., 2020. Effect of an E-learning module on personal protective equipment proficiency among prehospital personnel: Web-based randomized controlled trial. Journal of Medical Internet Research 22(8), e21265. DOI: 10.2196/21265.

12. Wu, S.H., Huang, C.C., Huang, S.S., Yang, Y.Y., Liu, C.W., Shulruf, B., Chen, C.H., 2020. Effect of virtual reality training to decreases rates of needle stick/sharp injuries in new-coming medical and nursing interns in Taiwan. Journal of Educational Evaluation for Health Professions 17, 1. DOI: 10.3352/jeehp.2020.17.1.