Educational Research Review 28 (2019) 100293

Contents lists available at ScienceDirect

Educational Research Review

journal homepage: www.elsevier.com/locate/edurev

Review

Webinars in higher education and professional training: A meta-analysis and systematic review of randomized controlled trials

Andreas Gegenfurtner[a,b,*], Christian Ebner[b]

[a] University of Passau, Germany
[b] Technische Hochschule Deggendorf, Germany

ARTICLE INFO

Keywords:
Webinar
web conferencing
Synchronous
Computer
Distance education
Student achievement
RCT
Best evidence synthesis

ABSTRACT

Digital learning environments are increasingly popular in higher education and professional training. Teaching and learning via webinars, and web conferencing more broadly, represents one widely used approach. Webinars are defined as web-based seminars, in which participants and facilitators communicate live over the Internet across distant geographical locations using shared virtual platforms and interact ubiquitously and synchronously in real time via voice over IP technology and web camera equipment. In the past, studies have reported mixed evidence concerning the effectiveness of webinars in promoting student achievement. As a remedy, this systematic literature review and meta-analysis cumulates observed effect sizes from previously published randomized controlled trials and corrects artifactual variance induced by sampling error. The research questions were: How effective are webinars in promoting student achievement? And which characteristics moderate webinar effectiveness? The findings suggest that webinars were slightly more effective than control conditions (online asynchronous learning management systems and offline face-to-face classroom instruction), but the differences were trivial in size. Differences were moderated by webinar, participant, achievement, and publication characteristics. This meta-analysis is the first to systematically review and meta-analyze the best evidence available for evaluating the effectiveness of webinars and video conferences in promoting student knowledge and skills. The implications of the study's findings can inform school teachers, lecturers, trainers, technologists, and theorists interested in the computer-supported design, implementation, delivery, tutoring, and assessment of webinar-based learning environments.

1. Introduction

How effective are webinars in promoting student achievement in higher education and professional training? And which characteristics moderate webinar effectiveness? The use of webinars and web conferencing systems in education has gained growing attention in recent years (Goe, Ipsen, & Bliss, 2018; Häkkinen & Järvelä, 2006; McKinney, 2017; McMahon-Howard & Reimers, 2013; Nelson, 2010; Olson & McCracken, 2015; Stout et al., 2012; Wang & Hsu, 2008), largely because webinars offer digital learning environments that students can access ubiquitously from anywhere with computer devices (Ebner & Gegenfurtner, 2019; Gegenfurtner, Zitt, & Ebner, in press; Tseng, Cheng, & Yeh, 2019). For example, Nicklen, Keating, Paynter, Storr, and Maloney (2016) examined webinar-based learning by physiotherapy students, and Harned et al. (2014) evaluated mental health webinars in the

* Corresponding author. University of Passau, Dr.-Hans-Kapfinger-Straße 14b, 94032, Passau, Germany.
E-mail address: andreas.gegenfurtner@uni-passau.de (A. Gegenfurtner).

https://doi.org/10.1016/j.edurev.2019.100293
Received 6 November 2018; Received in revised form 17 September 2019; Accepted 30 September 2019
Available online 07 October 2019
1747-938X/

Table 1
Definition of terms included in the meta-analytic review.

Term	Definition
Webinar	Web-based seminar, in which students and teachers are connected live across distant geographical locations using shared virtual platforms, such as Adobe Connect or Cisco WebEx, and interact synchronously in real time via voice over IP and web camera equipment
Webinar-based learning	Learning processes and outcomes in webinar contexts
Web conferencing	Umbrella term for different kinds of online meetings (educational and non-educational). A webinar is a special case of web conferencing that serves an educational function
Webcast	Web-based broadcasting over the Internet (live or on-demand) of audio or audiovisual information designed for single-to-many interaction, such as streams of online TV or radio or business presentations
Learning management system	Platform for online content that affords asynchronous interaction among students and teachers, such as Moodle

context of professional training. Webinars are frequently integrated into the curricula of distance education and blended learning programs (Cornelius & Gordon, 2013; Gegenfurtner, Schwab, & Ebner, 2018; Kear, Chetwynd, Williams, & Donelan, 2012; Khechine, Lakhal, Pascot, & Bytha, 2014; Testers, Gegenfurtner, Van Geel, & Brand-Gruwel, 2019; Wang & Hsu, 2008).

One problem with research on webinars and web conferencing in the educational technology literature is the typically small sample size. Individual study findings are therefore likely to be influenced by sampling error. This may explain some of the disagreements in the literature. In particular, some authors reported that webinar participants had higher learning outcomes than control participants (Alnabelsi, Al-Hussaini, & Owens, 2015; Kanter, Tsai, Holman, & Koerner, 2013; Spalla, 2012). Others reported findings in the opposite direction (Carrick et al., 2017; Constantine, 2012; Joshi, Thukral, Joshi, Deorari, & Vatsa, 2013). To account for effect size heterogeneity, the present study used meta-analytic methods to cumulate individual research findings on webinar effectiveness after controlling for sampling error. The goal was to synthesize the best available evidence from randomized controlled trials (RCTs). A second purpose was to estimate the extent to which moderator characteristics—features of the webinar, their participants, how achievement outcomes were assessed, or when and where the study was published—would moderate the extent of webinar effectiveness.

1.1. Webinars

The global trend of digitalization has also transformed the way in which education is designed, delivered, and implemented (Cook et al., 2008; Cuban, 2001; Margulieux, McCracken, & Catrambone, 2016; McKinney, 2017; Säljö, 2019; Siewiorek & Gegenfurtner, 2010; Siewiorek, Gegenfurtner, Lainema, Saarinen, & Lehtinen, 2013; Testers, Gegenfurtner, & Brand-Gruwel, 2015). Webinars are a common choice from the kaleidoscope of digital learning environments. Being an emerging field of research, however, the terminology is yet inexact. Table 1 offers some definitions to help enhance conceptual clarity. The term *webinar* is a neologism and portmanteau of the words *web* and *seminar*. In its simplest understanding, a webinar is a seminar that happens online over the Internet rather than offline in a traditional classroom. Like in all cases of *web conferencing*, communication among webinar participants (students and teachers) is mediated technologically via web cameras and voice over IP. Students and teachers can interact online from virtually anywhere worldwide; there is no need to travel to a physical seminar room. This ubiquity and geographical flexibility is an obvious advantage of webinars over traditional, offline face-to-face lectures. A webinar is considered a special case of web conferencing insofar as the function of a webinar is intrinsically educational in nature. While web conferences as an umbrella term can include, among others, meetings between business partners or video chats among friends or peers, webinars serve the purposes of learning and teaching. As such, webinars stand in contrast to *webcasts*, such as the streaming of online TV or radio or company presentations which are used for leisure, entertainment, or business but not necessarily for education. Interaction in webcasts is typically designed for single-to-many while interaction in webinars is typically designed for many-to-many. Furthermore, interaction in webinars is live, synchronous, and in real time, unlike interaction in *learning management systems* in which interaction is typically asynchronous (Ebner & Gegenfurtner, 2019; McKinney, 2017; Wang & Hsu, 2008). All of these digital environments are considered artifacts that afford and mediate the processes and practices of learning and teaching. As Säljö (2019, p. 27) puts it, "technologies are not just representations of the world, rather they are constitutive elements of the enactment of thinking and reasoning in social practices."

Typically, the timeline of a webinar starts with the planning phase which incorporates scheduling the webinar event and inviting participants to register online. Today, one webinar can technically host up to 3,000 participants, yet it seems likely that this number will expand in the future. Webinars with a smaller number of students are far more common, however (see, as examples, Alnabelsi et al., 2015; Carrick et al., 2017; Gegenfurtner et al., in press; Joshi et al., 2013; Kanter et al., 2013). For participants, the technical requirements include a fast internet connection as well as a browser or app installed on their digital device, such as a laptop, mobile phone, or tablet. A single teacher or a team of multiple teachers and/or technologists then prepare the virtual meeting room. This room is typically afforded within a web conferencing platform; examples include Adobe Connect and Cisco WebEx. During the webinar, features can include didactical or instructional activities that were also performed in a traditional, offline seminar. Typical online features afforded by contemporary webinar technology include screen sharing, video, slides, chats, Q&A, polls, virtual rooms for group work, and real-time feedback among students and teachers to facilitate webinar-based learning. At the end of a webinar event, teachers, facilitators, and technologists can perform follow-up analyses and evaluations of the webinar effectiveness.

1.2. Effectiveness of webinar-based learning

Students, tutors, and lecturers frequently report that they are satisfied with or enjoyed participating in webinar-based learning environments (Cornelius & Gordon, 2013; Gegenfurtner et al., 2018; Kear et al., 2012; Wang & Hsu, 2008). To date, however, no systematic review or meta-analysis has specifically focused on the effectiveness of webinar-based learning environments in promoting student achievement. Previous reviews covered, for example, blended learning (Liu et al., 2016; Margulieux et al., 2016; Means, Toyama, Murphy, & Baki, 2013), computer-supported collaborative learning (Gegenfurtner, Veermans, & Vauras, 2013), distance education (Bernard et al., 2004), online education (Means, Toyama, Murphy, Baki, & Jones, 2009), simulation-based learning (Gegenfurtner, Quesada-Pallarès, & Gegenfurtner, 2014), web-based learning (Cook, Garside, Levinson, Dupras, & Montori, 2010), different treatment interactions (Bernard et al., 2009), training methods in human resource development (Martin, Kolomitro, & Lam, 2014), or particular populations, including health care professionals (Richmond, Copsey, Hall, Davies, & Lamb, 2017), post-secondary students (Schmid, Bernard, Borokhovski, Tamim, Abrami, et al., 2016), or medical students (Cook et al., 2008; Taveira-Gomes, Ferreira, Taveira-Gomes, Severo, & Ferreira, 2016). Given that no systematic literature review or meta-analysis has yet targeted webinars, a meta-analytic review on webinar effectiveness in promoting student achievement seems timely.

When students participate in a webinar-based learning environment, the effectiveness of webinars can be assessed in several ways. First, it can be assessed in terms of participants' development from pretest to posttest, measuring their relative increase in knowledge and skills. For example, Alnabelsi et al. (2015) examined medical students' knowledge of otolaryngological emergencies before and after attending a webinar. In the present meta-analytic review, this first analysis is labeled the *PrePost* analysis of webinar effectiveness.

Second, webinar effectiveness can be assessed as the difference in achievement outcomes between webinar and control participants at posttest. For example, Spalla (2012) examined nursing students' intercultural competence at the end of a webinar intervention and compared their competence levels with a group of randomly assigned control participants. This second analysis is labeled the *WebinarControl* analysis.

Third, and arguably the most relevant for determining the effectiveness of webinars in promoting student achievement, we can compare how much webinar and control participants gained in knowledge and skills from pretest to posttest, taking into account their levels of prior knowledge before the intervention started. For example, Harned et al. (2014) randomly assigned therapy trainees to treatment conditions, measured their knowledge at baseline, and then estimated their relative gains in each condition. This third analysis is labeled the *Gain* analysis of webinar effectiveness. The major difference to the *WebinarControl* analysis is that the *Gain* analysis considers the level of prior knowledge before the intervention.

The meta-analytic review reported here compares webinar effectiveness on all three levels: PrePost, WebinarControl, and Gain. Effectiveness estimates are cumulated and synthesized from the best evidence reported in an RCT. In RCTs, participants are randomly assigned to treatment and control conditions. Studies that follow the RCT design limit sampling selection biases and are thus considered to offer the most robust scientific evidence, in terms of methodology (Schultz, Altman, Moher, & the CONSORT Group, 2010). Cheung and Slavin (2016) noted that effect sizes from randomized experiments are more conservative than quasi-experimental studies, which report higher mean effect sizes: "If quasi-experiments tend to overstate effect sizes, this implies that mean effect sizes from reviews that average randomized and quasi-experimental effect sizes are likely to be reporting inflated mean effect sizes" (p. 288). For these reasons (Cheung & Slavin, 2016; Schultz et al., 2010), the present meta-analytic review focuses on the best evidence and synthesizes effect sizes reported in RCTs to estimate how effective webinar-based learning environments are in promoting student achievement.

1.3. Boundary conditions of webinar effectiveness

Webinar effectiveness can be moderated by a number of boundary conditions. Consequently, we examined the cumulated effect size estimates as a function of webinar, participant, achievement, and publication characteristics, largely because these analyses extend the existing literature on webinar-based learning. Moreover, these analyses allow stronger inferences to be made as to whether effectiveness estimates are an artifact of certain boundary conditions or are indeed consistently identified across conditions. Comparing effect size estimates across subgroups can therefore contribute to the robustness of the present meta-analytic review, as it helps reveal patterns across moderating characteristics.

First, webinar characteristics include the length of webinar events and how often webinar events were cast. For example, some webinars are as short as 30 min (e.g., Nelson, 2010) while other webinars can last several hours (e.g., Spalla, 2012; Stout et al., 2012). Similarly, some webinars are single events (e.g., McMahon-Howard & Reimers, 2013) while other webinar events are repeated several times (e.g., Carrick et al., 2017). It seems likely that duration and frequency of webinars can moderate the effectiveness of webinar-based learning. In addition, the webinar technology (for example, Adobe Connect, Cisco WebEx, Elluminate, etc.) can moderate webinar effectiveness. Because the technological infrastructure of webinar enabling systems has probably improved over the years, it can be speculated that more recently cast webinars are more effective than webinars cast years ago. Finally, one can assume that instruction given during the webinar can moderate how and how much participants learn (Bernard et al., 2009; Gegenfurtner, Quesada-Pallarès, & Knogler, 2014; Means et al., 2013).

Second, participant characteristics can function as a boundary condition for webinar effectiveness. Specifically, in two Canadian samples, Khechine et al. (2014) and Lakhal, Khechine, and Pascot (2013) reported evidence that age and gender moderated the intentions of participants to use a webinar. This justifies consideration of the relative difference between female and male participants and among participants of varying ages in analyses of webinar effectiveness. Similarly, webinars are designed and implemented for differing populations of learners, including students in higher education (e.g., Olson & McCracken, 2015) and in professional training (e.g., Constantine, 2012); these populations tend to vary in terms of their ages, career trajectories, and work experience. Participant population is thus considered a boundary condition in the present meta-analytic review.

Third, achievement characteristics include the criteria by which achievement is defined and assessed. For example, achievement outcomes can be measured in terms of knowledge (Alnabelsi et al., 2015) on a multiple-choice test (Carrick et al., 2017) or as performance (Kanter et al., 2013) evaluated by trained raters (Stout et al., 2012). Thus, achievement type and assessment type are included as potential boundary conditions of webinar effectiveness (Gegenfurtner et al., 2013). More importantly, to estimate webinar effectiveness, the condition to which the webinars are compared matters. Some studies compared student achievement among webinar participants with control participants who received no training (McMahon-Howard & Reimers, 2013); other studies used participants who received traditional face-to-face instruction in classrooms as a control group (Joshi et al., 2013); and still other studies compared synchronous webinars with asynchronous learning management systems (Harned et al., 2014). It is likely that levels of webinar effectiveness will differ as a function of the type of control condition that is used.

Finally, in the meta-analytic literature, it is a well-known problem that effect sizes vary as a function of publication type: effects published in peer-reviewed journal articles tend to be more significant than those in unpublished sources or dissertations (Cheung & Slavin, 2016; Schmidt & Hunter, 2015; Siddaway, Wood, & Hedges, 2019). It is thus reasonable to compare effect sizes over different kinds of publication subgroups. In addition, as with webinar year, it can be assumed that publication year is a boundary condition of webinar effectiveness if it is true that webinar technology has improved over the years. Publication type and publication year are thus considered as two moderator variables in this study.

1.4. The present study

The present study focuses on the effectiveness of webinars in higher education and professional training. The dependent variable was student achievement, and the independent variable was participation in a webinar. Following an interest in offering evidence-based recommendations for webinar-based learning and education, the study sought to answer the research questions: In the context of higher education and professional training, how effective are webinars in promoting student achievement? And which characteristics moderate webinar effectiveness? The study pursued the dual goal of examining these two research questions.

The first goal was to cumulate observed effect sizes by correcting the variance across studies for the biases of unequal sample sizes and sampling error. We estimated the relative effectiveness of webinars in three ways: (a) by comparing pretest and posttest scores of webinar participants in a PrePost analysis; (b) by comparing posttest scores of webinar and control group participants in a WebinarControl analysis; and (c) by comparing the gain scores—the development from pretest to posttest—between webinar and control participants in a Gain analysis.

In addition to correcting these observed effect sizes, the second goal of the study was to estimate the extent to which a set of a priori hypothesized moderator variables might explain the remaining variance in results. Moderators were expected at four levels: webinar, participant, achievement, and publication. At the webinar level, we estimated the effects of webinar length, webinar year, webinar instruction, webinar technology, and whether there were single or repeated webinars. At the participant level, we examined the effects of age, gender, and population. At the achievement level, we were interested in the effects of achievement type, assessment type, and type of control group. Finally, at the publication level, we estimated the effects of publication year and publication type.

2. Methods

This meta-analysis and systematic review of webinars in higher education and professional training was performed in adherence to the standards of quality for conducting and reporting meta-analyses detailed in the PRISMA statement (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009), in the Methodological Quality Instrument (Tamim, Borokhovski, Bernard, Schmid, & Abrami, 2015), and in best practice guidelines (Beretvas, 2019; Cooper, 2016; Siddaway et al., 2019). Below, we report the inclusion and exclusion criteria, the literature search, the literature coding, interrater analyses, and the meta-analytic methods used.

2.1. Inclusion and exclusion criteria

Studies were identified that reported effect sizes on the relative effectiveness of webinars compared to control conditions in higher education or professional training. Table 2 presents the criteria for inclusion and exclusion in the meta-analysis. To be included, a study had to report a sample size for the webinar and control condition as well as mean and standard deviation estimates for both groups, or other effect sizes that could be converted to the standardized mean difference Cohen's d (e.g., md, F, t, X^2, or z statistics). Studies were included if they reported outcomes of randomized controlled trials; studies that did not randomly assign participants to the webinar and control conditions were omitted. Studies were also excluded if they focused only on asynchronous learning environments. Because the focus of inquiry was on the cognitive indicators of webinar effectiveness, we included studies that reported objectively measured knowledge or performance scores; studies using self-report data or studies focusing on non-cognitive outcomes were omitted. To minimize publication bias, we included original empirical research reported in peer-reviewed journal articles, book chapters, books, conference proceedings, and unpublished theses or dissertations, independent of publication language, disciplinary field, webinar technology, content, or instruction. We searched for studies published in a 15-year time frame, from January 2003 to December 2018. We used January 2003 as the starting point to continue and update the meta-analysis of Bernard et al. (2004), who ended their literature search in December 2002.

Table 2
Criteria for inclusion and exclusion in the meta-analytic review.

Criterion	Inclusion	Exclusion
Study design	Randomized controlled trials	Studies without randomization and/or without control condition
Synchrony	Synchronous webinars	Asynchronous environments
Effect sizes	Studies reporting an effect size and the sample size in both the webinar and any control groups	Studies not reporting an effect size or sample size information
Achievement type	Knowledge or performance scores	Other variables (e.g., attitudes)
Assessment type	Objective tests or ratings	Other sources (e.g., self-ratings)
Publication type	Original research, including peer-reviewed journal articles or unpublished dissertations	Literature reviews, meta-analyses, editorials, book reviews
Publication language	All languages	–
Publication date	January 2003–December 2018	Prior to 2003
Age group	Adult samples	Non-adult samples
Participant population	Higher education, professional training	Primary or secondary education
Webinar instruction	All instructional formats	–
Webinar technology	All technologies	–
Webinar content	All contents	–
Disciplinary field	All fields	–

An example of both discarded and retained articles may help illustrate how the inclusion criteria were applied. On the one hand, the randomized controlled trial by Goe et al. (2018), which compared synchronous webinar training with a no-training condition, was excluded because the dependent variable was not student achievement; rather, it was self-reported feelings of preparedness to help customers use social media tools for their career development. On the other hand, the study by Carrick et al. (2017) was included because the dependent variable, knowledge of neuro-otology, was measured by multiple-choice questions; the resulting test scores were then compared between participants who were randomly assigned to traditional face-to-face and online webinar classrooms.

2.2. Literature search

Using these inclusion criteria, the literature was searched in several ways. First, we performed an electronic search of four databases—ERIC, PsycINFO, PubMed, and Scopus—for any publication type in any language published between January 2003 and December 2018, using relevant keywords included in the title or abstract. The keywords were *webinar*, *webconference*, *webconferencing*, *web conference*, *web conferencing*, *web seminar*, *webseminar*, *adobeconnect*, *adobe connect*, *elluminate*, and *webex*. These terms were combined with *training*, *adult education*, *further education*, *continuing education*, and *higher education* to identify relevant articles on webinars in higher education and training.

This first search yielded a total of 616 hits: 101 from ERIC, 70 from PsycINFO, 129 from PubMed, and 316 from Scopus. Two trained raters checked all identified articles and eliminated 159 duplicates that were listed in more than one database; 457 articles remained. Both raters then screened a random subset of the identified 457 articles (10 percent; $k = 46$ articles) independently and in duplicate; as interrater agreement was high with Cohen's $\kappa = 0.93$ (95% CI = 0.79–1.00), one rater continued to screen the remaining studies for eligibility by reading titles and abstracts. A total of 406 hits were removed at this stage because the articles reported qualitative research, were review papers or commentaries, or focused on asynchronous learning management systems or modules.

The full texts of the remaining 51 articles were read in detail by both raters independently to check for eligibility. After the full texts were read, 44 records were excluded because the study was not a randomized controlled trial (29 removals), did not include a synchronous webinar condition (eleven removals), did not report sample size information separately for the webinar and the control condition (one removal[1]), or did not report empirical data (four removals). This left seven articles that were included in the meta-analysis.

In the second step of the literature search, we cross-referenced the included articles to identify other relevant studies. Cross-referencing included a backward search, in which we browsed through the lists of references in each article, and a forward search, in which we used Google Scholar to identify studies that cited the included articles. We also consulted the reference lists of fourteen earlier reviews and meta-analyses on online and distance education (Bernard et al., 2009; Cook et al., 2008, 2010; Gegenfurtner et al., 2013, 2014; Liu et al., 2016; Margulieux et al., 2016; Martin et al., 2014; McKinney, 2017; Means et al., 2009, 2013; Richmond et al., 2017; Schmid et al., 2014; Taveira-Gomes et al., 2016). This second step of the literature search yielded another five publications that met all inclusion criteria.

In summary, the seven articles obtained from the first search of electronic databases and the five articles obtained from the second search through cross-referencing resulted in a total set of twelve studies that included $k = 15$ independent data sources that were subsequently coded. Fig. 1 presents the study selection flowchart. Estimates of interrater agreement associated with the literature search are presented in section 2.4. All studies included in the meta-analysis are preceded by an asterisk in the list of references.

[1] We contacted the authors twice to obtain sample sizes, but did not receive the requested information. Sample size information is necessary for meta-analyses to calculate Hedges' g (Hedges, 1981) and to correct for sampling error (Schmidt & Hunter, 2015).

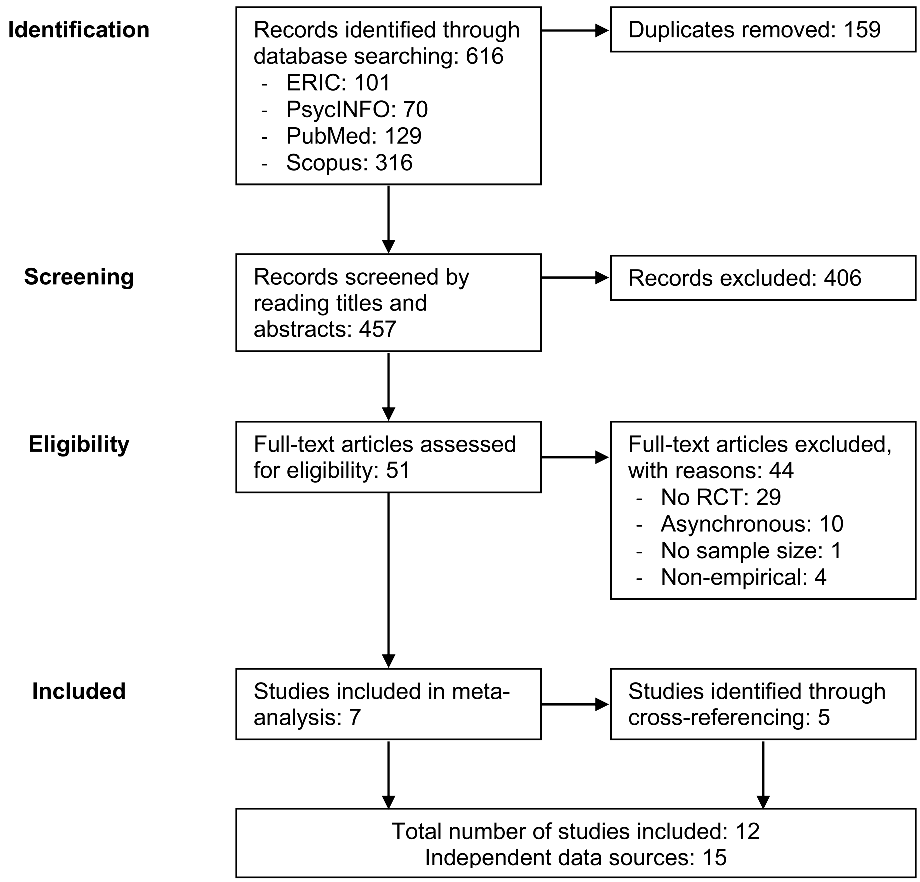

Fig. 1. Study selection flow diagram.

2.3. Literature coding

Once selected, two trained raters coded all studies independently and in duplicate using the coding scheme detailed in Table 3. Because one of the aims of this meta-analysis was to identify moderator variables that could account for effect size heterogeneity, several characteristics were tabulated from the research literature. Specifically, each study was coded for effect size estimates, achievement characteristics, webinar characteristics, participant characteristics, and publication characteristics.

The extracted effect sizes included the sample size in the webinar group and the control group at time 1 (pretest) and time 2 (posttest), the standardized mean difference Cohen's *d* between pretest to posttest in the webinar group ("PrePost"), the standardized mean difference Cohen's *d* of the posttest between the webinar group and the control group ("WebinarControl"), and the standardized mean difference Cohen's *d* of the gain scores from pretest to posttest between both groups ("Gain").

Webinar characteristics were coded as webinar duration (in hours), webinar instruction (lecture, lecture with interactive elements, interactive elements), the year when the webinar was cast, single or repeated webinar sessions, and webinar technology (Adobe Connect, Cisco WebEx, other).

Participant characteristics in the webinar and control conditions were coded as age (mean age in years), gender (percentage of females), and population (higher education, professional training).

Achievement characteristics were coded as achievement type (knowledge, performance) and assessment type (test, rating). In addition, we coded the control condition as asynchronous online training, synchronous face-to-face training, and no training.

Publication characteristics were coded as the first author, publication year, and publication type (peer-reviewed journal article, unpublished dissertation).

Some examples may help to illustrate coding decisions and how we applied the coding scheme to tabulate information from the individual studies. First, Harned et al. (2014) reported pretest (time1) and posttest (time 2) scores together with two follow-up measures six weeks (time 3) and twelve weeks (time 4) after training; to enable comparison with the other studies, the immediate posttest at time 2 was used to calculate Cohen's *d* estimates. Second, Olson and McCracken (2015) and Harned et al. (2014) compared an asynchronous online condition with a condition that combined asynchronous modules with synchronous webinars; in both studies, we used the combined condition as the measure for the webinar group because the primary difference in the two conditions was the additional webinar. A complete description of all coded study information is provided below in section 3.1.

Table 3
Coding scheme.

Main Category	Sub-Category
Publication characteristics	
Author	Name of first author
Publication year	Coded as year
Publication type	1 = peer-reviewed journal article
	2 = unpublished dissertation
	3 = other
Effect size estimates	
Sample size	Sample size *N* of the webinar group (pretest)
	Sample size *N* of the webinar group (posttest)
	Sample size *N* of the control group (pretest)
	Sample size *N* of the control group (posttest)
Effect size	Cohen's *d* of the pretest-posttest difference in the webinar group (web pre-post)
	Cohen's *d* of the posttest difference between the webinar and control group (web control)
	Cohen's *d* of the gain scores from pretest to posttest between both groups (gain)
Achievement characteristics	
Achievement type	1 = knowledge
	2 = performance
Assessment type	1 = objective test
	2 = rating
Control condition	1 = asynchronous online training
	2 = synchronous face-to-face training
	3 = no training
Webinar characteristics	
Webinar duration	Coded in hours
Webinar instruction	1 = lecture
	2 = lecture with interactive elements (case discussions, role plays, etc.)
	3 = interactive elements (case discussions, role plays, etc.)
Webinar year	Coded as year
Webinar technology	1 = Adobe Connect
	2 = Cisco WebEx
	3 = other
Single or repeated webinars	1 = single webinar
	2 = repeated webinars
Participant characteristics	
Participant age	Coded in years for participants in the webinar group
	Coded in years for participants in the control group
Participant gender	Percentage of female participants in the webinar group
	Percentage of female participants in the control group
Participant population	1 = higher education
	2 = professional training

2.4. Interrater reliability and agreement

Interrater reliability and agreement are important characteristics of any meta-analysis to ensure the credibility of the literature search and coding (Beretvas, 2019; Moher et al., 2009; Schmidt & Hunter, 2015; Tamin et al., 2015). Following the recommendations of Hoyt (2019), we used an intraclass correlation coefficient (ICC) to determine interrater reliability for continuous scales and Cohen's kappa coefficient (κ) to determine interrater agreement for nominal scales. ICC and κ provide robust estimates when there are, as in our case, exactly two raters who identified, screened, and coded the studies independently and in duplicate (Hoyt, 2019). Conflicts were resolved through consensus. ICC estimates and their 95% confidence intervals were calculated using SPSS 24 based on a mean-rating (k = 2), absolute-agreement, 2-way mixed-effects model (Koo & Li, 2016). Cohen's κ estimates and their standard errors were calculated using SPSS 24, and the standard errors were used to compute the 95% confidence intervals around κ. Table 4 presents all

Table 4
Interrater reliability and agreement of the literature search and coding.

Step	Category	Coefficient	Estimate	95% Confidence Interval
Literature Search	Identification	κ	0.99	0.98–1.00
	Screening	κ	0.93	0.79–1.00
	Eligibility	κ	1.00	1.00–1.00
Literature Coding	Continuous	ICC	0.97	0.95–0.98
	Categorical	κ	0.95	0.90–1.00

Note. κ = Cohen's kappa coefficient; ICC = intraclass correlation coefficient.

estimates for interrater agreement and interrater reliability for the literature search and coding.

To judge the relative strength of interrater agreement, Landis and Koch (1977) proposed to classify κ values between 0.41 and 0.60 as moderate agreement, values between 0.61 and 0.80 as substantial agreement, and values between 0.81 and 1.00 as almost perfect agreement. To judge the relative strength of interrater reliability, Koo and Li (2016) suggested to classify ICC values between 0.50 and 0.75 as moderate reliability, values between 0.75 and 0.90 as good reliability, and values greater than 0.90 as excellent reliability. Note that these threshold values for κ and ICC should be interpreted as rule-of-thumb benchmarks, not as definite values of rating quality (Koo & Li, 2016; Landis & Koch, 1977).

For the literature search, agreement was calculated separately for the identification, screening, and eligibility checks of studies, as detailed in the study selection flow diagram in Fig. 1. The respective κ estimates, with their 95% confidence intervals, indicate how reliable our decisions were to include or exclude studies in each step. The results were: for identification, $\kappa = 0.99$; for screening, $\kappa = 0.93$; and for eligibility, $\kappa = 1.00$, with the respective 95% confidence intervals ranging from 0.79 to 1.00. These values can be interpreted as substantial to almost perfect agreement (Landis & Koch, 1977).

For the literature coding, ICC estimates were calculated separately for the continuous and the categorical scales, as detailed in the coding scheme in Table 3. The continuous scales were publication year, all effect size estimates, webinar duration, group size, webinar year, participant age, and gender. The remaining scales were categorical. The respective κ and ICC estimates, with their 95% confidence intervals, indicated how reliably we extracted study information from the included articles. The results, ICC = 0.97 for the continuous scales and $\kappa = 0.95$ for the categorical scales, can be interpreted as being in almost perfect agreement (Landis & Koch, 1977) and having excellent reliability (Koo & Li, 2016).

2.5. Meta-analytic calculations

The meta-analytic calculations occurred in two stages. A primary meta-analysis aimed to compute the corrected effect size estimates. A meta-analytic moderator estimation then aimed to identify the a priori defined moderator effects in the corrected effect sizes. Both stages are specified below.

The primary meta-analysis was done following the procedures of meta-analysis of experimental effects (Schmidt & Hunter, 2015). We calculated Cohen's d estimates (a) for the difference between pretest and posttest measures in the webinar group (*PrePost*), using the formula $d = (M_{\text{Web post}} - M_{\text{Web pre}})/SD_{\text{pooled}}$; (b) for the difference in the posttest measures between the webinar and control groups (*WebinarControl*), using the formula $d = (M_{\text{Web post}} - M_{\text{Con post}})/SD_{\text{pooled}}$; and (c) for the difference in the mean pre–post gains in the webinar and control groups (*Gain*), using the formula $d = (M_{\text{g-e}} - M_{\text{g-c}})/SD_{\text{pooled-pre}}$ as detailed in Schmidt and Hunter (2015). When mean and standard deviation estimates were not available, we used formulae provided by Wan, Wang, Liu, and Tong (2014) to estimate M and SD from sample size, median, and range, and converted F into Cohen's d using formulae provided by Polanin and Snilstveit (2016). All Cohen's d estimates were transformed into Hedges' g (Hedges, 1981) and then into the point-biserial correlation coefficient r_{pb} (Schmidt & Hunter, 2015). Next, the distribution of r_{pb} was corrected for the attenuation effect of unequal sample sizes to get r_{us}; it was then further corrected for sampling error to get r_c using the formulae provided by Schmidt and Hunter (2015). Note that the correction was conducted using a weighted average, not Fisher's z transformation, since the latter was shown to produce upwardly biased correlation estimates (Hall & Brannick, 2002). Finally, we converted r_c into the sampling-error corrected Cohen's d_c (Polanin & Snilstveit, 2016) and computed standard errors and 95% confidence intervals of d_c.

The meta-analytic moderator estimation followed the primary meta-analysis and was used to estimate the effects of continuous and categorical moderator variables. Continuous moderator effects were estimated using an unrestricted weighted least squares (WLS) meta-regression analysis (Stanley & Doucouliagos, 2017). Categorical moderator effects were estimated using theory-driven subgroup analyses[2] (Schmidt & Hunter, 2015).

All calculations were based on the assumption that the population parameter value δ varies from study to study, so we used a random-effects model to obtain realistic estimates of the width of the confidence intervals (Cafri, Kromrey, & Brannick, 2010). Following the recommendations of Higgins and Green (2011), publication bias was analyzed in two ways: through visual inspection of funnel plots and through statistical estimation of funnel plot asymmetry using the Egger test (Egger, Smith, Schneider, & Minder, 1997). In addition, some authors advocate performing an outlier analysis. Because current approaches for identifying outlier coefficients in meta-analytic data sets tend to over-identify small correlations relative to large ones (Beal, Corey, & Dunlap, 2002), the removal of outlying cases becomes problematic. In addition, the formula for sampling error variance used in the present study allows and corrects for occasional extreme outlying values; it follows that eliminating outliers can overcorrect for sampling error and underestimate SE_{dc} (Schmidt & Hunter, 2015). Based on these problems with outlier removal in meta-analytic work, calculations in the present study were based on the full data set.

3. Results

3.1. Description of included studies

The twelve studies that were included in the analysis offered a total of 15 independent data sources with 36 effect sizes. The total

[2] A critique of using subgroup analysis is that it reduces the number of data sources per analysis, resulting in second-order sampling error; the possibility of second-order sampling error is therefore indicated when warranted for discussing the results (Schmidt & Hunter, 2015).

Table 5
Number of data sources, participants, and participant characteristics.

	Webinar Pretest	Webinar Posttest	Control Pretest	Control Posttest
k	11	15	11	15
Total N	591	716	571	698
Average N	53.73 (± 37.33)	47.73 (± 39.78)	51.91 (± 36.89)	46.53 (± 39.09)
Age	40.12 (± 2.83)	39.13 (± 2.48)	40.48 (± 4.45)	38.60 (± 2.95)
Gender	57.83 (± 0.22)	52.00 (± 0.19)	57.83 (± 0.21)	52.40 (± 0.18)

Note. Means (± standard deviations). k number of data sources, N sample size, age in years, gender as percentage of female participants.

sample size was 591 participants ($k = 11$) at pretest and 716 participants ($k = 15$) at posttest for the webinar condition; for the control condition, the total sample size was 571 participants ($k = 11$) at pretest and 698 participants ($k = 15$) at posttest. On average, studies had 53 (47) participants at pretest (posttest) in the webinar condition and 51 (46) participants at pretest (posttest) in the control condition; this small sample size in original studies signals the presence of sampling error, which tends to justify meta-analytic synthesis to correct for sampling error. Table 5 presents information on the number of data sources, participants, and participant characteristics for the webinar and control conditions at pre- and posttest. Differences in age and gender between conditions and measurement times were statistically nonsignificant ($p > 0.05$). Note that demographic information on age and gender was only sporadically reported in the original studies, with six studies reporting gender distributions and four studies reporting participant age. Because of this limited data pool and the associated risk of second-order sampling error (Schmidt & Hunter, 2015), age and gender were excluded from further analyses.

Table 6 presents a detailed description of all included studies, together with the coding of participant, achievement, and webinar characteristics. Concerning participants, seven datasets (46.67 percent) examined students in higher education and eight (53.33 percent) examined trainees in professional training. Concerning achievement type, ten datasets (66.67 percent) examined knowledge and five (33.33 percent) examined performance. Concerning assessment type, nine datasets (60.00 percent) used a standardized test and six used trained raters. Concerning the control condition, five datasets (33.33 percent) compared webinars with an asynchronous learning environment, seven (46.67 percent) with face-to-face teaching in classrooms, and three (20.00 percent) with a no-training condition. The total length of webinar instruction ranged from half an hour to 25 h ($M = 5.30$, $SD = 7.51$). In six datasets (40.00 percent), the webinar was a single event and not followed up, while in nine datasets (60 percent), webinars were repeated more than once. The webinar technology was Adobe Connect in four (26.67 percent), Cisco WebEx in six (40.00 percent), and another technology or not reported in five datasets (33.33 percent). Concerning publication type, four datasets (26.67 percent) were from unpublished doctoral dissertations and eleven (73.33 percent) were from peer-reviewed journal articles. Webinar year ranged from 2008 to 2012 (six datasets did not report when the webinar was cast) and publication year ranged from 2010 to 2017.

The effect sizes reported in the individual studies varied. In the PrePost analysis, Hedges' g ranged from 0.85 to 4.55, with a mean g of 2.14 ($SD = 1.11$). In the WebinarControl analysis, Hedges' g ranged from -0.36 to 0.68, with a mean g of 0.13 ($SD = 0.36$). In the Gain analysis, Hedges' g ranged from -0.28 to 1.13, with a mean g of 0.33 ($SD = 0.52$). This heterogeneity in effect sizes tends to warrant meta-analytic synthesis. Figs. 2–4 present the forest plot for the PrePost, Control, and Gain datasets.

3.2. Publication bias

Publication bias was analyzed visually and statistically. Visually, we inspected the funnel plots of each of the three datasets, PrePost, Control, and Gain, as shown in Figs. 5–7. Statistically, we estimated funnel plot asymmetry with the Egger test (Egger et al., 1997), which calculates a linear regression of the effect size against its standard error, weighted by the inverse of the variance of the effect size. The results of the Egger tests indicated a nonsignificant funnel plot asymmetry, with $\beta = 3.271$, $SE = 4.208$, $p = 0.457$; $\beta = 0.787$, $SE = 0.865$, $p = 0.381$; and $\beta = 3.561$, $SE = 1.750$, $p = 0.072$ for the analyses PrePost, WebinarControl, and Gain, respectively. In addition to these visual and statistical analyses, the literature search was deliberately wide and included all studies independent of publication type, publication language, disciplinary field, webinar technology, content, or instruction (see Table 2 for the inclusion criteria). In summary, based on the applied search criteria and the outcomes of the visual and statistical analyses, the presence of publication bias was unlikely in the current meta-analysis.

3.3. Results of primary meta-analyses

The psychometric properties of the primary meta-analyses are shown in Table 7. Of particular interest are the corrected Cohen's d estimates. Cohen (1988), Hyde (2005), and Rosenthal (1996) proposed to classify d values of 0.10 as trivial, values of 0.20 as small, values of 0.50 as medium, values of 0.80 as large, and values of 1.30 as very large. Findings of the $k = 11$ data sources ($N = 1{,}158$ participants) in the PrePost analysis show a d_c of 1.556 ($SE_{dc} = 0.067$, 95% $CI = 1.425$; 1.687), which indicates a very large effect of webinar participation on student achievement. Findings of the $k = 14$ data sources ($N = 1{,}291$ participants) in the WebinarControl analysis show a d_c of 0.140 ($SE_{dc} = 0.056$, 95% $CI = 0.030$; 0.250), which indicates a trivial-to-small effect on posttest student achievement in webinars compared to a control condition. Finally, findings of the $k = 11$ data sources ($N = 1{,}118$ participants) in the Gain analysis show a d_c of 0.151 ($SE_{dc} = 0.060$, 95% $CI = 0.033$; 0.269), which indicates a trivial-to-small effect on gains in student

Table 6
Description of participant, achievement, and webinar characteristics of the included studies.

	Participant characteristics			Achievement characteristics				
Study	Sample	Population		Achievement type		Assessment type		Control condition
		Information	Code	Information	Code	Information	Code	Information
Alnabelsi et al. (2015)	Medical students	Higher education	1	Knowledge of otolaryngological emergencies	1	Knowledge test	1	Face-to-face in lecture theatre
Carrick et al. (2017)	Healthcare professionals	Professional training	2	Knowledge of neuro-otology	1	Multiple choice test	1	Face-to-face in classroom
Constantine, 2012 (a)	Community health aides/practitioners	Professional training	2	Knowledge of clinical imaging	1	Knowledge test	1	Asynchronous computer-based training
Constantine, 2012 (b)	Community health aides/practitioners	Professional training	2	Performance in telehealth cases	2	Rubric score	2	Asynchronous computer-based training
Harned et al., 2014 (a)	Mental health professionals	Professional training	2	Knowledge of exposure therapy	1	Multiple choice test	1	Asynchronous online environment
Harned et al., 2014 (b)	Mental health professionals	Professional training	2	Clinical proficiency	2	Rating of role play performance	2	Asynchronous online environment
Joshi et al., 2013 (a)	Nursing students	Higher education	1	Knowledge of newborn care	1	Knowledge test	1	Self-reading in face-to-face classroom
Joshi et al., 2013 (b)	Nursing students	Higher education	1	OSCE score	2	OSCE	2	Self-reading in face-to-face classroom
Kanter et al. (2013)	Psychotherapists	Professional training	2	Vignette score of hypothetical therapy situations	1	Rating of vignette response	2	No training
McMahon-Howard (2013)	Child protective service employees	Professional training	2	Knowledge of commercial sexual exploitation of children	1	Knowledge test	1	No training
Nelson (2010)	Nursing students	Higher education	1	Knowledge of acid-base balance	1	Multiple choice test	1	Face-to-face in classroom
Nicklen et al. (2016)	Physiotherapy students	Higher education	1	Knowledge of physiotherapy during pregnancy	1	Multiple choice test	1	Face-to-face in classroom
Olson and McCracken (2015)	Undergraduate students	Higher education	1	Course grade	1	Not reported	–	Asynchronous online modules
Spalla (2012)	Nursing students	Higher education	1	Intercultural competence	1	Competence test	1	Face-to-face classroom
Stout et al. (2012)	Pediatricians	Professional training	2	Spirometry quality	2	Rating of spirometry sessions	2	No training

(continued on next page)

Table 6 (*continued*)

Study	Achievement characteristics	Webinar characteristics							
	Control condition	Webinar duration	Webinar technology		Webinar year	Webinar Instruction		Group size in webinar	Single or repeated webinars
	Code		Information	Code		Information	Code		
Alnabelsi et al. (2015)	2	Not reported	Cisco WebEx	2	Not reported	Lecture	1	25	Single
Carrick et al. (2017)	2	25 h	Adobe Connect	1	Not reported	Problem-solving exercises and discussions (flipped classroom)	3	Not reported	Repeated
Constantine, 2012 (a)	1	0,75 h	Adobe Connect	1	2011	Lecture, polls, videos	2	Not reported	Single
Constantine, 2012 (b)	1	0,75 h	Adobe Connect	1	2011	Lecture, polls, videos	2	Not reported	Single
Harned et al., 2014 (a)	1	8 h	Cisco WebEx	2	2012	Discussions, role plays, active practice of core concepts	2	8	Repeated
Harned et al., 2014 (b)	1	8 h	Cisco WebEx	2	2012	Discussions, role plays, active practice of core concepts	2	8	Repeated
Joshi et al., 2013 (a)	2	Not reported	Not reported	–	2011	Lecutre, discussions, drills, demonstrations	2	28	Repeated
Joshi et al., 2013 (b)	2	Not reported	Not reported	–	2011	Lecutre, discussions, drills, demonstrations	2	28	Repeated
Kanter et al. (2013)	3	Not reported	iCohere	3	Not reported	Personal exercises, didactic presentatios, case discussions	2	8	Repeated
McMahon-Howard (2013)	3	1,5 h	Not reported	–	2012	Lecture	1	Not reported	Single
Nelson (2010)	2	0,5 h	Cisco WebEx	2	2010	Lecture	1	Not reported	Single
Nicklen et al. (2016)	2	1,5 h	Cisco WebEx	2	2012	Case-based learning	2	5	Single
Olson and McCracken (2015)	1	Not reported	Adobe Connect	1	Not reported	Lecture with text-based chatting	2	Not reported	Repeated
Spalla (2012)	2	2 h	GoogleChat	3	2012	Discussion sessions	3	Not reported	Repeated
Stout et al. (2012)	3	5 h	Cisco WebEx	2	2008	CD-ROM tutorial, interactive case-based sessions	2	Not reported	Repeated

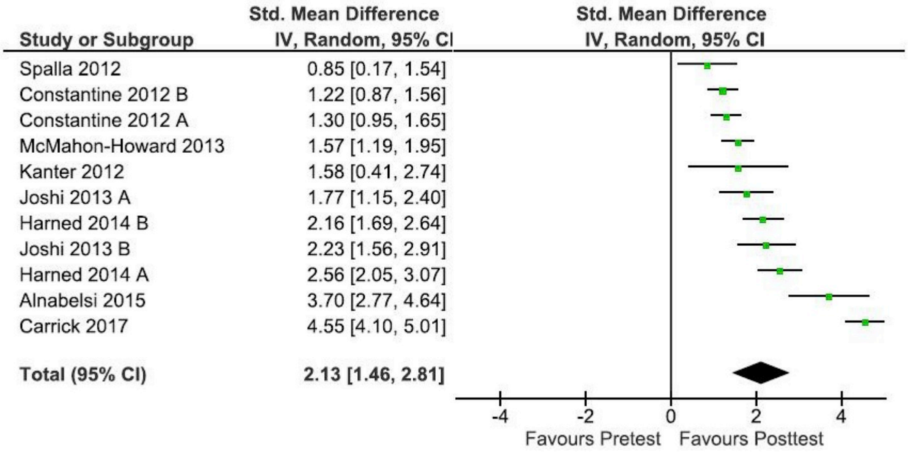

Fig. 2. Forest plot of the PrePost analysis.

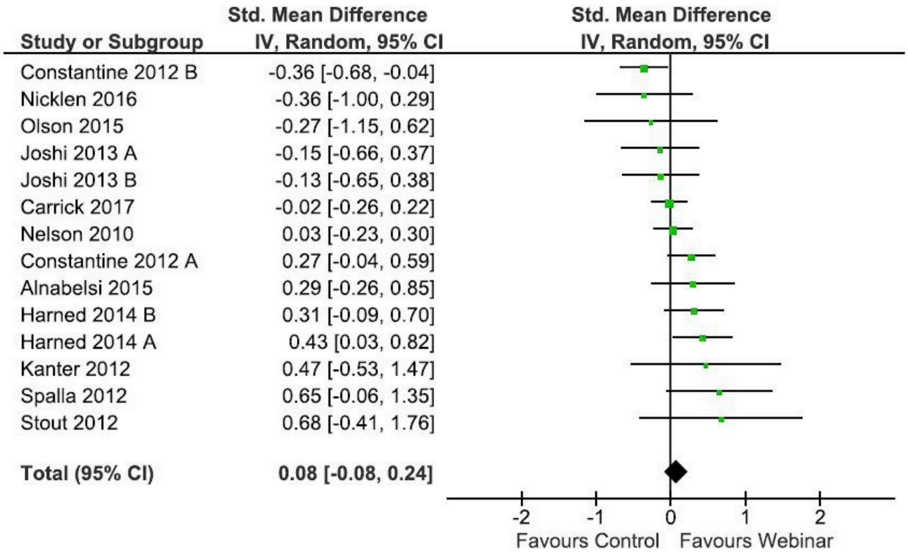

Fig. 3. Forest plot of the WebinarControl analysis.

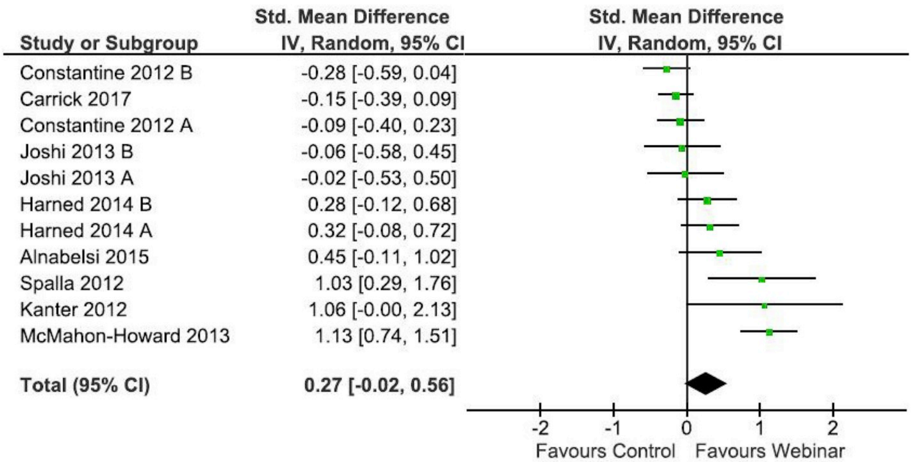

Fig. 4. Forest plot of the gain analysis.

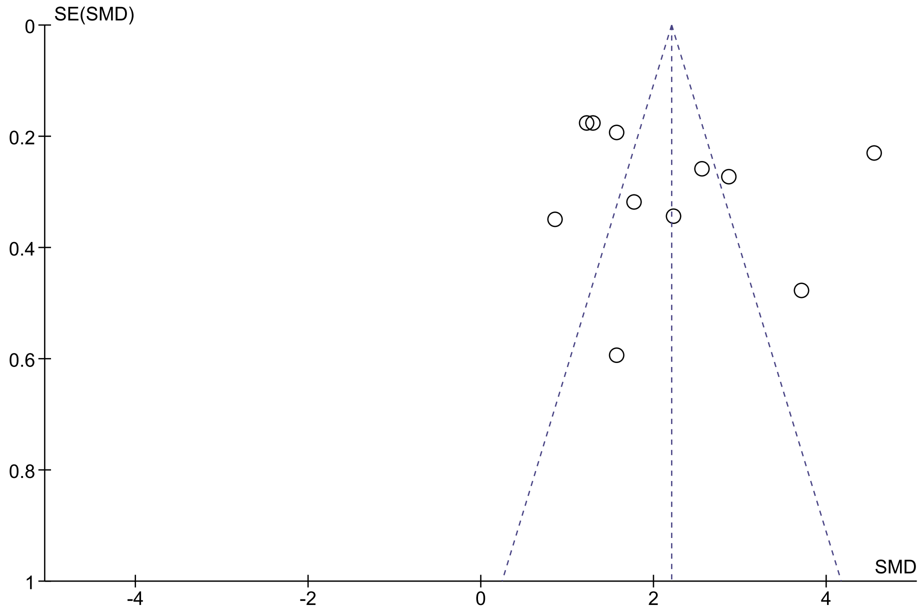

Fig. 5. Funnel plot of the PrePost analysis.

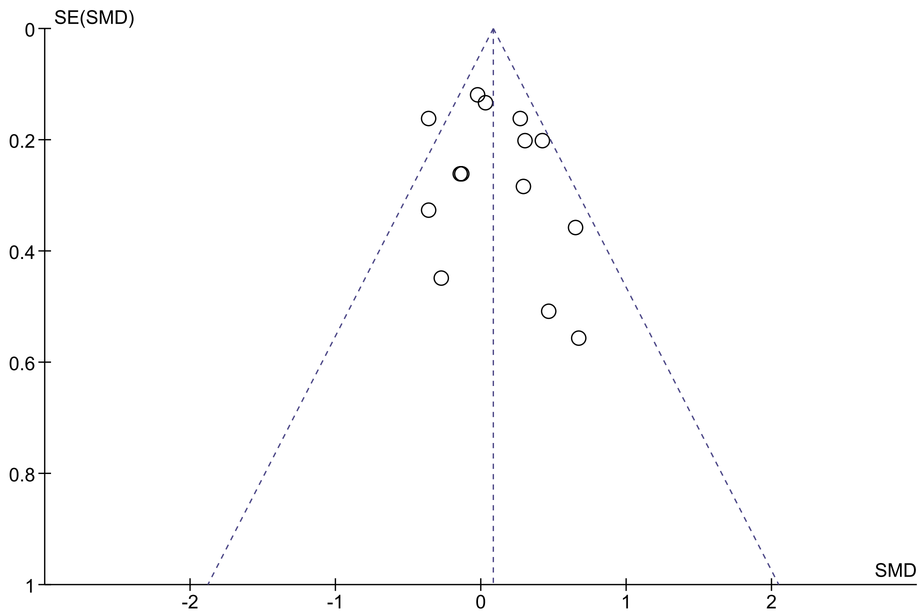

Fig. 6. Funnel plot of the WebinarControl analysis.

achievement in webinars compared to a control condition. All effect sizes are in the positive direction, suggesting that webinar participants benefitted from online synchronous education.

3.4. Results of meta-analytic moderator estimation

Table 8 presents the findings of the meta-analytic moderator estimation of the PrePost analysis for the categorical moderator variables. The findings indicate that peer-reviewed journal articles reported significantly higher achievement scores from pretest to posttest than unpublished dissertations, X^2 (1) = 17.50, $p < 0.001$. Webinar technology was another statistically significant moderator, with studies using Adobe Connect reporting lower effect sizes than studies using other technologies, X^2 (3) = 11.69, $p = 0.009$. Publication year produced a statistically significant moderator effect, $\beta = 0.990$, $p < 0.001$, with stronger increases from pretest to posttest in more recently published studies. Webinar duration was another statistically significant moderator, $\beta = 0.997$, $p < 0.001$, with stronger increases from pretest to posttest in longer webinars.

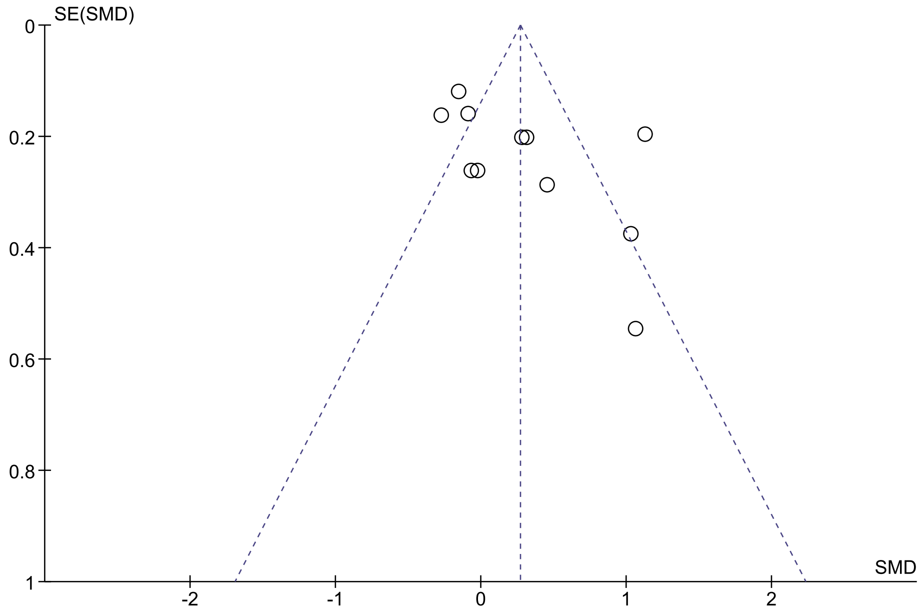

Fig. 7. Funnel plot of the gain analysis.

Table 7
Psychometric properties of the primary meta-analyses.

	k	N	d	g	r_{pb}	r_{us}	r_c	d_c	SE_{dc}	95% CI
PrePost	11	1,158	2.166	2.135	0.675	0.707	0.614	1.556	0.067	1.425; 1.687
WebinarControl	14	1,291	0.137	0.131	0.063	0.132	0.070	0.140	0.056	0.030; 0.250
Gain	11	1,118	0.343	0.335	0.149	0.224	0.075	0.151	0.060	0.033; 0.269

Note. k = number of data sources, N = total sample size; d = Cohen's d; g = Hedges' g; r_{pb} = pointbiserial correlation coefficient; r_{us} = r_{pb} corrected for the attenuation effect of unequal sample sizes; r_c = r_{us} corrected for sampling error; d_c = r_c converted to d; SE_{dc} = standard error of d_c; *95% CI* = 95% confidence interval around d_c.

Table 8
Psychometric properties of the meta-analytic moderator estimation for the pre-post analysis.

Moderator	k	N	d	g	r_{pb}	r_{us}	r_c	d_c	SE_{dc}	*95% CI*	X^2
Achievement											X^2 (1) = 1.18, p = 0.28
Knowledge	7	822	2.354	2.329	0.686	0.695	0.626	1.605	0.080	1.448; 1.762	
Performance	4	336	1.835	1.798	0.655	0.727	0.582	1.431	0.123	1.190; 1.672	
Technology											X^2 (3) = 11.69, p = 0.009
Adobe Connect	3	582	2.367	2.357	0.660	0.466	0.502	1.161	0.090	0.985; 1.134	
Cisco WebEx	3	270	2.843	2.807	0.801	0.784	0.744	2.227	0.155	1.923; 2.531	
Other	2	52	1.270	1.215	0.506	0.872	0.830	2.976	0.410	2.172; 3.780	
Not specified	3	254	1.883	1.857	0.676	0.760	0.687	1.891	0.152	1.593; 2.189	
Assessment Source											X^2 (1) = 0.39, p = 0.53
Objective Test	6	786	2.602	2.575	0.735	0.684	0.620	1.580	0.082	1.419; 1.741	
Rating	5	372	1.642	1.608	0.602	0.734	0.600	1.500	0.118	1.269; 1.731	
Publication Type											X^2 (1) = 17.50, p < 0.001
Dissertation	3	344	1.133	1.123	0.485	0.507	0.420	0.926	0.114	0.703; 1.149	
Empirical Journal	8	814	2.553	2.515	0.746	0.781	0.695	1.933	0.085	1.766; 2.010	
Context											X^2 (1) = 0.82, p = 0.37
Higher Education	4	198	2.175	2.138	0.670	0.867	0.874	3.597	0.231	3.144; 4.050	
Professional Training	7	960	2.160	2.134	0.677	0.615	0.560	1.352	0.072	1.211; 1.493	
Instruction											–
Lecture	1	142	–	–	–	–	–	–	–	–	
Lecture + Interactive	9	742	1.963	1.930	0.654	0.733	0.618	1.572	0.084	1.407; 1.737	
Interactive Elements	1	274	–	–	–	–	–	–	–	–	
Repeated											X^2 (1) = 2.42, p = 0.12
Single Webinar	4	500	1.968	1.948	0.641	0.566	0.485	1.109	0.096	0.921; 1.297	
Repeated Webinars	7	658	2.279	2.243	0.694	0.787	0.711	2.022	0.096	1.834: 2.210	

Table 9
Psychometric properties of the meta-analytic moderator estimation for the WebinarControl analysis.

Moderator	k	N	d	g	r_{pb}	r_{us}	r_c	d_c	SE_{dc}	95% CI	X^2
Achievement											X^2 (1) = 0.06, $p = 0.80$
Knowledge	9	950	0.099	0.097	0.046	0.019	0.039	0.078	0.065	−0.049; 0.205	
Performance	5	341	0.206	0.194	0.092	0.335	0.157	0.318	0.109	0.104; 0.532	
Control Group											X^2 (2) = 3.48, $p = 0.18$
Asynchronous	5	526	0.074	0.076	0.037	−0.045	0.108	0.217	0.088	0.045; 0.389	
Face-to-Face	7	735	0.047	0.044	0.020	0.045	0.010	0.020	0.074	−0.125; 0.165	
No Training	2	30	0.610	0.575	0.275	0.876	0.872	3.563	0.609	2.369; 4.757	
Technology											X^2 (3) = 6.13, $p = 0.11$
Adobe Connect	4	602	−0.098	−0.095	−0.048	−0.149	0.033	0.066	0.082	−0.095; 0.227	
Cisco WebEx	6	524	0.238	0.230	0.111	0.192	0.092	0.185	0.088	0.013; 0.357	
Other	2	49	0.585	0.560	0.269	0.765	0.743	2.220	0.371	1.493; 2.947	
Not specified	2	116	−0.145	−0.140	−0.070	−0.120	−0.120	−0.242	0.188	−0.610; 0.126	
Assessment Source											X^2 (1) = 0.11, $p = 0.74$
Objective Test	8	917	0.028	0.028	0.013	−0.067	0.015	0.030	0.066	−0.099; 0.159	
Rating	6	374	0.283	0.270	0.128	0.396	0.205	0.419	0.105	0.213; 0.625	
Publication Type											X^2 (1) = 0.01, $p = 0.93$
Dissertation	4	565	0.153	0.148	0.070	0.169	0.100	0.201	0.085	0.034; 0.368	
Empirical Journal	10	726	0.131	0.125	0.060	0.117	0.047	0.094	0.074	−0.051; 0.239	
Context											X^2 (1) = 0.71, $p = 0.40$
Higher Education	7	481	0.010	0.009	0.002	−0.034	−0.006	−0.012	0.091	−0.190; 0.166	
Professional Training	7	810	0.264	0.254	0.123	0.297	0.123	0.248	0.071	0.109; 0.387	
Instruction											–
Lecture	1	224	–	–	–	–	–	–	–	–	
Lecture + Interactive	12	793	0.159	0.153	0.073	0.153	0.114	0.229	0.071	0.090; 0.368	
Interactive Elements	1	274	–	–	–	–	–	–	–	–	
Repeated											X^2 (1) = 0.90, $p = 0.34$
Single Webinar	5	620	−0.024	−0.026	−0.013	−0.035	0.050	0.100	0.081	−0.059; 0.256	
Repeated Webinars	9	671	0.227	0.219	0.104	0.224	0.089	0.179	0.077	0.028; 0.330	

Table 10
Psychometric properties of the meta-analytic moderator estimation for the gain analysis.

Moderator	k	N	d	g	r_{pb}	r_{us}	r_c	d_c	SE_{dc}	95% CI	X^2
Achievement											X^2 (1) = 0.16, $p = 0.69$
Knowledge	7	791	0.387	0.381	0.171	0.216	0.091	0.183	0.071	0.044; 0.322	
Performance	4	327	0.266	0.253	0.112	0.239	0.038	0.076	0.111	−0.142; 0.294	
Control Group											X^2 (2) = 23.50, $p < 0.001$
Asynchronous	4	506	0.058	0.060	0.030	0.045	0.022	0.044	0.089	−0.130; 0.218	
Face-to-Face	5	473	0.256	0.250	0.113	0.231	0.078	0.156	0.092	−0.024; 0.336	
No Training	2	139	1.130	1.095	0.479	0.564	0.258	0.534	0.174	0.193; 0.875	
Technology											X^2 (3) = 17.36, $p < 0.001$
Adobe Connect	3	582	−0.172	−0.170	−0.083	−0.049	−0.044	−0.088	0.083	−0.251; 0.075	
Cisco WebEx	3	248	0.353	0.350	0.172	0.237	0.202	0.413	0.129	0.160; 0.666	
Other	2	49	1.085	1.045	0.463	0.900	0.880	3.705	0.481	2.762; 4.648	
Not specified	3	239	0.353	0.350	0.150	0.034	0.069	0.138	0.130	−0.117; 0.393	
Assessment Source											X^2 (1) = 0.52, $p = 0.47$
Objective Test	6	758	0.277	0.273	0.123	0.111	0.058	0.116	0.073	−0.027; 0.253	
Rating	5	360	0.423	0.408	0.181	0.360	0.112	0.225	0.106	0.017; 0.433	
Publication Type											X^2 (1) = 1.19, $p = 0.28$
Dissertation	3	341	0.227	0.223	0.094	0.241	0.028	0.056	0.109	−0.158; 0.270	
Empirical Journal	8	777	0.386	0.376	0.170	0.218	0.096	0.193	0.072	0.052; 0.334	
Context											X^2 (1) = 0.57, $p = 0.45$
Higher Education	4	199	0.358	0.350	0.159	0.296	0.223	0.458	0.144	0.176; 0.740	
Professional Training	7	919	0.335	0.326	0.144	0.183	0.043	0.086	0.066	−0.043; 0.215	
Instruction											–
Lecture	1	123	–	–	–	–	–	–	–	–	
Lecture + Interactive	9	721	0.309	0.300	0.136	0.258	0.098	0.197	0.075	0.050; 0.344	
Interactive Elements	1	274	–	–	–	–	–	–	–	–	
Repeated											X^2 (1) = 0.30, $p = 0.59$
Single Webinar	4	481	0.308	0.305	0.134	0.115	0.047	0.094	0.096	−0.094; 0.282	
Repeated Webinars	7	637	0.363	0.351	0.158	0.286	0.096	0.193	0.080	0.036; 0.350	

Table 9 presents the findings of the meta-analytic moderator estimation of the WebinarControl analysis for the categorical moderator variables. Publication year produced a statistically significant moderator effect, $\beta = -0.862$, $p < 0.001$, with smaller posttest differences in more recently published studies. Studies with a control group that received no training ($d_c = 3.563$) reported higher effect sizes than studies with control groups that received asynchronous ($d_c = 0.217$) or face-to-face instruction ($d_c = 0.020$). Studies using other webinar technologies ($d_c = 2.220$) reported higher effect sizes than studies using Adobe Connect ($d_c = 0.066$), Cisco WebEx ($d_c = 0.185$), or unspecified technologies ($d_c = -0.242$).

Table 10 presents the findings of the meta-analytic moderator estimation of the Gain analysis for the categorical moderator variables. Studies with a control group that received no training ($d_c = 0.534$) reported higher effect sizes than studies with control groups that received asynchronous ($d_c = 0.044$) or face-to-face instruction ($d_c = 0.156$). Webinar technology was another statistically significant moderator, with studies using Adobe Connect reporting lower effect sizes than studies using other technologies, X^2 (3) = 17.36, $p < 0.001$. The effect of webinar year was on the edge of significance, $\beta = 0.794$, $p = 0.059$, favoring webinars cast more recently. But because six datasets did not report when the webinar was cast, this effect must be interpreted with caution.

4. Discussion

This meta-analysis and systematic literature review aimed to offer a synthesis of the best available evidence from randomized controlled trials, examining the effectiveness of webinars and web conferencing to promote student achievement in higher education and professional training. In this section, we summarize the main findings of the effectiveness of webinars and what moderates such effectiveness, the practical relevance of the findings for educational technologists who design and implement webinar-based learning environments, and limitations and future research directions that follow from the presented meta-analytic evidence.

4.1. Main findings

The first research question of this meta-analytic review concerned the effectiveness of webinars and web conferencing in promoting student achievement. The development of webinar participants from pretest to posttest showed a very large effect, with a corrected Cohen's d estimate of 1.556. This is, per se, a positive finding. But more interesting than a comparison of participant knowledge before and after webinars is to compare their knowledge and skills with people who participated in other learning environments, taking into account any levels of prior knowledge. This comparison was performed with the Gain analysis. Do people learn more in webinars than elsewhere? When we compared how strongly participants in webinars and participants in other environments developed their knowledge and skills, and when we considered their prior knowledge before the intervention, then we got a corrected Cohen's d estimate of 0.151. This was a positive effect that favored webinars; however, the effect was trivial in size. To answer the first research question, webinars are trivially more effective in promoting student achievement than other learning environments.

The second research question of this meta-analytic review was concerned with the boundary conditions of webinar effectiveness. Meta-analytic moderator estimation suggested that gain scores were larger when webinar participants were compared to control participants who received no training at all ($d_c = 0.534$). When compared with students who were randomly assigned to asynchronous learning management systems ($d_c = 0.044$) or traditional face-to-face classrooms ($d_c = 0.156$), differences in Gain scores were positive but trivial in size. The comparison of the two most widely used web conferencing technologies showed that WebEx ($d_c = 0.413$) tended to be more effective than Adobe Connect ($d_c = -0.088$); however, due to the small cell sizes and the associated likelihood of second-order sampling error, this effect size difference should not be interpreted as a causal effect of one technology being more effective than another.

4.2. Practical relevance

What are the implications of the findings for educational technologists? How can instructional designers and practitioners use the available meta-analytic evidence when creating and implementing webinar-based learning environments? The practical relevance of this study is associated with three aspects.

First, participants developed more knowledge and skills from pretest to posttest when webinar duration was longer. This positive association between duration and learning seems intuitive because participants had longer exposure and thus more time to develop their knowledge and skills. Interestingly, and contrary to our assumptions, repeated webinar events were not more effective than single webinar events. The presented evidence did not inform about an “optimal” webinar length, so educational technologists and practitioners can continue to rely on experience-based rules of thumb. Still, we can conclude that longer webinar durations are positively associated with more knowledge and skills.

Second, educational technologists must often choose the kind of webinar technology they want to implement. Based on the meta-analytic evidence presented in the moderator analysis, Cisco WebEx is positively associated with participants' gain scores. But as noted previously, the cell sizes and the respective sample sizes were small; this indicates the likelihood of second-order sampling error. Furthermore, from the perspective of instructional design, how a technology is used and implemented is more important for the learner than which technology is used (Bernard et al., 2009; Gegenfurtner et al., 2014).

Third, many of the examined moderator variables were statistically nonsignificant. Hence, web conferencing seems to be equally effective for learners in higher education and professional training, for webinars that focus on declarative knowledge or procedural skills, when learning is assessed with knowledge tests or performance ratings, and when webinar events are sequenced over multiple occasions or cast as single events. Unfortunately, there were too few studies that systematically varied the instructional design approach within webinars. We assume that differences in webinar effectiveness will emerge when different interaction treatments are compared (see also Bernard et al., 2009; Gegenfurtner et al., 2014; Wang & Hsu, 2008).

Finally, this meta-analytic review shows that webinars and face-to-face classroom teaching are comparable in their effectiveness to promote student learning. This is good news for all teachers, trainers, and lecturers who wish to offer and implement digital modes of learning for their students because webinars offer higher levels of flexibility for the learners even if achievement effects of webinars are small (positive, but trivial in size). This is because students can attend lectures at home or at their workplace without the temporal and monetary cost of traveling (Cornelius & Gordon, 2013; Gegenfurtner et al., 2018; Gegenfurtner et al., in press; Lakhal et al., 2013; Wang & Hsu, 2008).

4.3. Limitations and directions for future research

This study has some limitations that should be noted. The first limitation is that the study meta-analytically corrected for the artifactual variance induced by sampling error and the bias of unequal sample sizes in the experimental and control groups. However, the original research reports may be affected by additional biases, such as extraneous factors introduced by study procedure (Schmidt & Hunter, 2015). To lessen this bias, we performed a series of meta-analytic moderator analyses. Still, the true score population estimates of webinar effectiveness may be somewhat greater than those reported in this study.

A further limitation is that some cell sizes for some of the moderating conditions were small, as were the associated sample sizes. However, some meta-analysts have noted that correcting for bias at a small scale mitigates sampling error compared to uncorrected estimates in individual studies (Schmidt & Hunter, 2015). Still, although most of the cells contained sample sizes in the hundreds, some did contain fewer; this indicates underestimation of sampling error and the likelihood of second-order sampling error in those few cases.

Finally, the study considered a total of twelve moderator variables. It was implicitly assumed that the a priori selected achievement, webinar, participant, and publication characteristics were among the most dominant sources for effect size heterogeneity in webinar effectiveness (McKinney, 2017; Wang & Hsu, 2008). Nonetheless, the total number of potential moderator variables likely exceeds twelve. Results of the meta-analytic moderator analysis are therefore limited in their generalizability across the full range of possible boundary conditions. It is worth noting that a majority of studies was situated within the medical and health sciences, possibly because the methodological standards of randomized controlled trials are more frequently applied in clinical settings. Once a larger pool of original studies exists, it would be worth examining potential domain contingencies as a moderating boundary condition.

Directions for future research include extending the base of primary studies to broadening the possibilities of meta-analytic synthesis. In particular, further research can aim to systematically vary the instructional approach within and across webinars to estimate the extent to which different designs of interactive treatment can promote (or hinder) gains in knowledge and skills of webinar participants. In the present meta-analytic review, the cell sizes for the moderator variable “webinar instruction” were too small to warrant quantitative comparison; the only remedy for this is additional original empirical investigations to broaden and extend the meta-analytic database. Systematic manipulation of features has shown to produce differential effects on learning and achievement in a number of digital learning scenarios (for reviews, see e.g. Bernard et al., 2009; Gegenfurtner et al., 2013; Gegenfurtner et al., 2014; Means et al., 2013); thus, following these lines of research, it seems likely that future studies manipulating instructional approaches during webinars can produce empirical evidence on how to further improve webinar effectiveness.

5. Conclusion

As noted at the outset, the interest in webinars and web conferencing has grown considerably in the educational technology literature (Carrick et al., 2017; Gegenfurtner et al., 2018; Gegenfurtner et al., in press; Goe et al., 2018; McKinney, 2017; Tseng et al., 2019; Wang & Hsu, 2008). While much of this interest has shown, on an anecdotal level, that webinars work well for students in higher education and professional training, this study sought to answer empirically how effective webinars are, under which boundary conditions, using the best evidence reported in randomized controlled trials. Results of our meta-analysis and systematic literature review of 15 independent data sources with 36 effect sizes, comparing 716 webinar participants with 698 control participants under twelve moderating conditions, indicated that webinars were positively associated with gains in knowledge and skills. Future research is encouraged to extend the analyses reported here to the examination of webinar effectiveness under varying interaction treatments.

Acknowledgements

This work was funded funded by the German Bundsministerium für Bildung und Forschung/ce:grant-sponsor > (BMBF; grant number: 16OH22004).

References

*Alnabelsi, T., Al-Hussaini, A., & Owens, D. (2015). Comparison of traditional face-to-face teaching with synchronous e-learning in otolaryngology emergencies teaching to medical undergraduates: A randomised controlled trial. *European Archives of Oto-Rhino-Laryngology, 272*, 759–763. https://doi.org/10.1007/s00405-

014-3326-6.
Beal, D. J., Corey, D. M., & Dunlap, W. P. (2002). On the bias of Huffcutt and Arthur's (1995) procedure for identifying outliers in the meta-analysis of correlations. *Journal of Applied Psychology, 87*, 583–589. https://doi.org/10.1037/0021-9010.87.3.583.
Beretvas, S. N. (2019). Meta-analysis. In G. R. Hancock, L. M. Stapleton, & R. O. Mueller (Eds.). *The reviewer's guide to quantitative methods in the social sciences* (pp. 260–268). (2nd ed.). New York: Routledge.
Bernard, R. M., Abrami, P. C., Borokhovski, E., Wade, C. A., Tamim, R. M., Surkes, M. A., et al. (2009). A meta-analysis of three types of interaction treatments in distance education. *Review of Educational Research, 79*, 1243–1289. https://doi.org/10.3102/0034654309333844.
Bernard, R. M., Abrami, P. C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., et al. (2004). How does distance education compare to classroom instruction? A meta-analysis of the empirical literature. *Review of Educational Research, 74*, 379–439. https://doi.org/10.3102/00346543074003379.
Cafri, G., Kromrey, J. D., & Brannick, M. T. (2010). A meta-meta-analysis: Empirical review of statistical power, type I error rates, effect sizes, and model selection of meta-analyses published in psychology. *Multivariate Behavioral Research, 45*, 239–270. https://doi.org/10.1080/00273171003680187.
*Carrick, F. R., Abdulrahman, M., Hankir, A., Zayaruzny, M., Najem, K., Lungchukiet, P., et al. (2017). Randomized controlled study of a remote flipped classroom neuro-otology curriculum. *Frontiers in Human Neuroscience, 8*, 349. https://doi.org/10.3389/fneur.2017.00349.
Cheung, A. C. K., & Slavin, R. E. (2016). How methodological features affect effect sizes in education. *Educational Researcher, 45*, 283–292. https://doi.org/10.3102/0013189X1665.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
*Constantine, M. B. (2012). *A study of individual learning styles and e-learning preferences among community health aides/practitioners in rural Alaska.* Cypress, CA: Unpublished doctoral dissertation, Trident University International.
Cook, D. A., Garside, S., Levinson, A. J., Dupras, D. M., & Montori, V. M. (2010). What do we mean by web-based learning? A systematic review of the variability of interventions. *Medical Education, 44*, 765–774. https://doi.org/10.1111/j.1365-2923.2010.03723.x.
Cook, D. A., Levinson, A. J., Carside, S., Dupras, D. M., Erwin, P. J., & Montori, V. M. (2008). Internet-based learning in the health professions: A meta-analysis. *Journal of the American Medical Association, 300*, 1181–1196. https://doi.org/10.1001/jama.300.10.1181.
Cooper, H. (2016). *Research synthesis and meta-analysis* (5th ed.). Thousand Oaks, CA: Sage.
Cornelius, S., & Gordon, C. (2013). Facilitating learning with web conferencing recommendations based on learners' experiences. *Education and Information Technologies, 18*, 275–285. https://doi.org/10.1007/s10639-012-9241-9.
Cuba, L. (2001). *Oversold and underused: Computers in the classroom.* Cambridge, MA: Harvard University Press.
Ebner, C., & Gegenfurtner, A. (2019). Learning and satisfaction in webinar, online, and face-to-face instruction: A meta-analysis. *Frontiers in Education, 4*, 92. https://doi.org/10.3389/feduc.2019.00092.
Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*, 629–634. https://doi.org/10.1136/bmj.315.7109.629.
Gegenfurtner, A., Quesada-Pallarès, C., & Knogler, M. (2014). Digital simulation-based training: A meta-analysis. *British Journal of Educational Technology, 45*, 1097–1114. https://doi.org/10.1111/bjet.12188.
Gegenfurtner, A., Schwab, N., & Ebner, C. (2018). "There's no need to drive from A to B": Exploring the lived experience of students and lecturers with digital learning in higher education. *Bavarian Journal Applied Sciences, 4*, 310–322. https://doi.org/10.25929/bjas.v4i1.50.
Gegenfurtner, A., Veermans, K., & Vauras, M. (2013). Effects of computer support, collaboration, and time lag on performance self-efficacy and transfer of training: A longitudinal meta-analysis. *Educational Research Review, 8*, 75–89. https://doi.org/10.1016/j.edurev.2012.04.001.
Gegenfurtner, A., Zitt, A., & Ebner, C. (in press). Evaluating webinar-based training: A mixed methods study on trainee reactions toward digital web conferencing. International Journal of Training and Development.
Goe, R., Ipsen, C., & Bliss, S. (2018). Pilot testing a digital career literacy training for vocational rehabilitation professionals. *Rehabilitation Counseling Bulletin, 61*, 236–243. https://doi.org/10.1177/0034355217724341.
Häkkinen, P., & Järvelä, S. (2006). Sharing and constructing perspectives in web-based conferencing. *Computers & Education, 47*, 433–447. https://doi.org/10.1016/j.compedu.2004.10.015.
Hall, S. M., & Brannick, M. T. (2002). Comparison of two random-effects methods of meta-analysis. *Journal of Applied Psychology, 87*, 377–389. https://doi.org/10.1037/0021-9010.87.2.377.
*Harned, M. S., Dimeff, L. A., Woodcook, E. A., Kelly, T., Zavertnik, J., Contreras, I., et al. (2014). Exposing clinicians to exposure: A randomized controlled dissemination trial of exposure therapy for anxiety disorders. *Behavior Therapy, 45*, 731–744. https://doi.org/10.1016/j.beth.2014.04.005.
Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*, 107–128. https://doi.org/10.2307/1164588.
Higgins, J. P. T., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions (version 5.1.0). The Cochrane Collaboration*. Available from: www.handbook.cochrane.org.
Hoyt, W. T. (2019). Interrater reliability and agreement. In G. R. Hancock, L. M. Stapleton, & R. O. Mueller (Eds.). *The reviewer's guide to quantitative methods in the social sciences* (pp. 132–144). (2nd ed.). New York: Routledge.
Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60*, 581–592. https://doi.org/10.1037/0003-066X.60.6.581.
*Joshi, P., Thukral, A., Joshi, M., Deorari, A. K., & Vatsa, M. (2013). Comparing the effectiveness of webinars and participatory learning on essential newborn care (ENBC) in the class room in terms of acquisition of knowledge and skills of student nurses: A randomized controlled trial. *Indian Journal of Pediatrics, 80*, 168–170. https://doi.org/10.1007/s12098-012-0742-8.
*Kanter, J. W., Tsai, M., Holman, G., & Koerner, K. (2013). Preliminary data from a randomized pilot study of web-based functional analytic psychotherapy therapist training. *Psychotherapy, 50*, 248–255. https://doi.org/10.1037/a0029814.
Kear, K., Chetwynd, Williams, J., & Donelan, H. (2012). Web conferencing for synchronous online tutorials: Perspectives of tutors using a new medium. *Computers & Education, 58*, 953–963. https://doi.org/10.1016/j.compedu.2011.10.015.
Khechine, H., Lakhal, S., Pascot, D., & Bytha, A. (2014). UTAUT model for blended learning: The role of gender and age in the intention to use webinars. *Interdisciplinary Journal of E-Learning and Learning Objects, 10*, 33–52. https://doi.org/10.28945/1994.
Koo, T. K., & Li, M. Y. (2016). A guideline for selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*, 155–163. https://doi.org/10.1016/j.jcm.2016.02.012.
Lakhal, S., Khechine, H., & Pascot, D. (2013). Student behavioural intentions to use desktop video conferencing in a distance course: Integration of autonomy to the UTAUT model. *Journal of Computing in Higher Education, 25*, 93–121. https://doi.org/10.1007/s12528-013-9069-3.
Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174. https://doi.org/10.2307/2529310.
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. https://doi.org/10.2196/jmir.4807.
Margulieux, L. E., McCracken, W. M., & Catrambone, R. (2016). A taxonomy to define courses that mix face-to-face and online learning. *Educational Research Review, 19*, 104–118. https://doi.org/10.1016/j.edurev.2016.07.001.
Martin, B. O., Kolomitro, K., & Lam, T. C. M. (2014). Training methods: A review and analysis. *Human Resource Development Review, 13*, 11–35. https://doi.org/10.1177/1534484313497947.
McKinney, W. P. (2017). Assessing the evidence for the educational efficacy of webinars and related internet-based instruction. *Pedagogy Health Promotion: The Scholarship of Teaching and Learning, 3*, 475–515. https://doi.org/10.1177/2373379917700876.
*McMahon-Howard, J., & Reimers, B. (2013). An evaluation of a child welfare training program on the commercial sexual exploitation of children (CSEC). *Evaluation and Program Planning, 40*, 1–9. https://doi.org/10.1016/j.evalprogplan.2013.04.002.
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.

Means, B., Toyama, Y., Murphy, R., Baki, M., & Jones, K. (2009). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies.* Washington, DC: US Department of Education.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *British Medical Journal, 339*, b2535. https://doi.org/10.1136/bmj.b2535.

*Nelson, L. S. (2010). *Learning outcomes of webinar versus classroom instruction among baccalaureate nursing students: A randomized controlled trial.* Denton, TX: Unpublished doctoral dissertation, Texas Woman's University.

*Nicklen, P., Keating, J. L., Paynter, S., Storr, M., & Maloney, S. (2016). Remote-online case-based learning: A comparison of remote-online and face-to-face, case-based learning – a randomized controlled trial. *Education and Health, 29*, 195–202. https://doi.org/10.4103/1357-6283.204213.

*Olson, J. S., & McCracken, F. E. (2015). Is it worth the effort? The impact of incorporating synchronous lectures into an online course. *Online Learning Journal, 19*, 73–84. https://doi.org/10.24059/olj.v19i2.499.

Polanin, J. R., & Snilstveit, B. (2016). *Campbell methods policy note on converting between effect sizes.* Oslo: The Campbell Collaboration.

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*, 227. https://doi.org/10.1186/s12909-017-1047.

Rosenthal, J. A. (1996). Qualitative descriptors of strength of association and effect size. *Journal of Social Service Research, 21*, 37–59. https://doi.org/10.1300/J079v21n04_02.

Säljö, R. (2019). Materiality, learning, and cognitive practices: Artifacts as instruments of thinking. In T. Cerrato-Pargman, & I. Jahnke (Eds.). *Emergent practices and material conditions in learning and teaching with technologies* (pp. 21–32). Cham: Springer.

Schmid, R. F., Bernard, R. M., Borokhovski, E., Tamim, R. M., Abrami, P. C., Surkes, M. A., et al. (2014). The effects of technology use in postsecondary education: A meta-analysis of classroom applications. *Computers & Education, 72*, 271–291. https://doi.org/10.1016/j.compedu.2013.11.002.

Schmidt, F. L., & Hunter, J. E. (2015). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd ed.). Thousand Oaks, CA: Sage.

Schultz, K. F., Altman, D. G., & Moher, D. the CONSORT Group. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *BMC Medicine, 8*, 18. https://doi.org/10.1186/1741-7015-8-18.

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology, 70*, 747–770. https://doi.org/10.1146/annurev-psych-010418-102803.

Siewiorek, A., & Gegenfurtner, A. (2010). Leading to win: The influence of leadership style on team performance during a computer game training. In K. Gomez, L. Lyons, & J. Radinsky (Vol. Eds.), *Learning in the disciplines: Vol. 1*, (pp. 524–531). Chicago, IL: ISLS.

Siewiorek, A., Gegenfurtner, A., Lainema, T., Saarinen, E., & Lehtinen, E. (2013). The effects of computer-simulation game training on participants' opinions on leadership styles. *British Journal of Educational Technology, 44*, 1012–1035. https://doi.org/10.1111/bjet.12084.

*Spalla, T. L. (2012). *Building the ARC in nursing education: Cross-cultural experiential learning enabled by the technology of video or web conferencing.* Columbus, OH: Unpublished doctoral dissertation, Ohio State University.

Stanley, T. D., & Doucouliagos, H. (2017). Neither fixed nor random: Weighted least squares meta-regression. *Research Synthesis Methods, 8*, 19–42. https://doi.org/10.1002/jrsm.1211.

*Stout, J. W., Smith, K., Zhou, C., Solomon, C., Dozor, A. J., Garrison, M. M., et al. (2012). Learning from a distance: Effectiveness of online spirometry training in improving asthma care. *Academic Pediatrics, 12*, 88–95. https://doi.org/10.1016/j.acap.2011.11.006.

Tamim, R. M., Borokhovski, E., Bernard, R. M., Schmid, R. F., & Abrami, P. C. (2015, April). A methodological quality tool for meta-analysis: The case of the educational technology literature. *Paper presented at the annual meeting of the American educational research association, Chicago, IL*.

Taveira-Gomes, T., Ferreira, P., Taveira-Gomes, I., Severo, M., & Ferreira, M. A. (2016). What are we looking for in computer-based learning interventions in medical education? A systematic review. *Journal of Medical Internet Research, 18*. https://doi.org/10.2196/jmir.5461 e204.

Testers, L., Gegenfurtner, A., & Brand-Gruwel, S. (2015). Motivation to transfer learning to multiple contexts. In L. Das, S. Brand-Gruwel, K. Kok, & J. Walhout (Eds.). *The school library rocks: Living it, learning it, loving it* (pp. 473–487). Heerlen: IASL.

Testers, L., Gegenfurtner, A., Van Geel, R., & Brand-Gruwel, S. (2019). From monocontextual to multicontextual transfer: Organizational determinants of the intention to transfer generic information literacy competences to multiple contexts. *Frontline Learning Research, 7*, 23–42. https://doi.org/10.14786/flr.v7i1.359.

Tseng, J.-J., Cheng, Y.-S., & Yeh, H.-N. (2019). How pre-service English teachers enact TPACK in the context of web-conferencing teaching: A design thinking approach. *Computers & Education, 128*, 171–182. https://doi.org/10.1016/j.compedu.2018.09.022.

Wan, X., Wang, W., Liu, J., & Tong, T. (2014). Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Medical Research Methodology, 14*, 135. https://doi.org/10.1186/1471-2288-14-135.

Wang, S.-K., & Hsu, H.-Y. (2008). Use of the webinar tool (Elluminate) to support training: The effects of webinar-learning implementation from student-trainers' perspective. *The Journal of Interactive Online Learning, 7*, 175–194.