Review of Educational Research
Month 202X, Vol. XX, No. X, pp. 1–33
DOI: 10.3102/0034654321990713
Article reuse guidelines: sagepub.com/journals-permissions

Video Improves Learning in Higher Education: A Systematic Review

Michael Noetel and Shantell Griffith
Australian Catholic University

Oscar Delaney
University of Queensland

Taren Sanders, Philip Parker, Borja del Pozo Cruz, and Chris Lonsdale
Australian Catholic University

Universities around the world are incorporating online learning, often relying on videos (asynchronous multimedia). We systematically reviewed the effects of video on learning in higher education. We searched five databases using 27 keywords to find randomized trials that measured the learning effects of video among college students. We conducted full-text screening, data extraction, and risk of bias in duplicate. We calculated pooled effect sizes using multilevel random-effects meta-analysis. Searches retrieved 9,677 unique records. After screening 329 full texts, 105 met inclusion criteria, with a pooled sample of 7,776 students. Swapping video for existing teaching methods led to small improvements in student learning (g = 0.28). Adding video to existing teaching led to strong learning benefits (g = 0.80). Although results may be subject to some experimental and publication biases, they suggest that videos are unlikely to be detrimental and usually improve student learning.

Keywords: multimedia, online learning, instructional design, cognitive load, active learning

In early 2020, most universities around the world scrambled to move all their classes online in response to the COVID-19 pandemic (Lee, 2020). Some observers fear that students receive a substandard experience online and that universities should transition back to in-person teaching as soon as possible (Deming, 2020). Others suggest that many students lack both social and technical support that are needed to access online learning (Warschauer & Matuchniak, 2010). This problem is exacerbated in less affluent countries (Valenzuela-Levi, 2020) and may create a harmful feedback loop where differential access may increase gaps in achievement (Johnson & Mejia, 2014). Although many such concerns may be valid, meta-analyses suggest that online and distance learning can be viable replacements for face-to-face teaching (Means et al., 2009, 2010; Means et al., 2013). Among the most common strategies for transitioning to online learning are synchronous videoconferences (e.g., Zoom, Adobe Connect) and asynchronous videos (henceforth, "video"; e.g., lecture recordings; Cook et al., 2010; Lee, 2020; Veletsianos & Houlden, 2020). Systematic reviews have shown that videoconferences are satisfactory substitutes for traditional pedagogies, with comparable learning outcomes (Chipps et al., 2012; Tomlinson et al., 2013). This review aims to identify the effects of video on student learning in higher education.

According to the cognitive theory of multimedia learning, videos, face-to-face classes, and videoconferences could all maximize the use of our cognitive infrastructure (Mayer, 2008). Our minds have distinct but connected neurological systems for processing auditory and visual information. Some media present content in one format or another: for example, books present visual information; podcasts and radio present auditory information. These systems are connected, such that a visual activity, like reading, will activate the auditory systems as we "hear" the words in our head. When listening to a podcast, our visual systems may construct a "picture" of what is being described. However, the cognitive theory of multimedia learning proposes that learning is more effective when complementary information is presented to both systems (Mayer, 2008). Meta-analyses of the multimedia effect confirm this proposition, showing that people learn better when both channels are used rather than just one or the other (Mayer, 2008; Rolfe & Gray, 2011). Video serves this purpose, communicating information through both channels. We define videos as prerecorded multimedia that combine moving images (pictures/graphics) and audio (usually spoken words or background sounds; Mayer, 2009; Mayer et al., 2020). The definition includes narrated animations or so-called 'voice-over-PowerPoint' but not presentation slides or podcasts alone. As mentioned earlier, video is not the only form of instruction that capitalizes on the multimedia effect by presenting complementary information to both audio and visual systems. High-quality face-to-face instruction and videoconferences (a.k.a., "Zoom," "Skype") can offer this experience, while also allowing for dynamic interaction between staff and students.

As a result, many different *media*—videos, lectures, videoconferences—can leverage these *methods* (e.g., using video and audio channels to communicate information). This distinction between media and method has a long and controversial history (Clark, 1983, 1994; Kozma, 1994; Warnick & Burbules, 2007). On one hand, many of the mechanisms by which media improve learning can be replicated in other forms of teaching, as noted above (Clark, 1983, 1994). For example, while a video can show a close-up of an authentic surgical operation, so can still images in a lecture, dummy surgeries in tutorials, or surgical internships. These *instructional methods*—in this case, authentic demonstrations—are often confounded with *instructional media*, and some argue that the media pales in importance compared with the method (Clark, 1983, 1994; Warnick & Burbules, 2007). On the other hand, some media are more capable of implementing successful instructional methods (Kozma, 1994). For example, computer-aided instruction is not necessary for giving each individual student personalized feedback and the right level of challenge—great teachers do this—but instructional technologies do make this kind of personalization easier. The media facilitates the method (Kozma, 1994). In the same way, there are a number of reasons why video might facilitate useful instructional methods.

If there are learning advantages for videos compared with these other approaches, they might be due to the control available to the learner or the editing available to the teacher. By being asynchronous, videos generally afford the learner more control over their learning, which may offer a series of benefits. First of all, perceived control can improve student motivation and regulate their cognitive load (Abeysekera & Dawson, 2015; S. Schneider, Nebel, et al., 2018). In higher education, motivational interventions are often successful when they aim to increase student autonomy and self-direction (Aelterman et al., 2019; Lazowski & Hulleman, 2016). In addition to motivational benefits, this sense of choice and control has been shown to mitigate real and perceived cognitive load (S. Schneider, Nebel, et al., 2018). Specifically, videos allow for students to manage their own cognitive load by pausing to take notes, rewinding difficult sections, or accelerating easy ones. Many academics can relate to being in a videoconference that they wished they could fast-forward. Video grants that capacity. The cost of producing quality video has rapidly declined as anyone with a smartphone can shoot and edit high-definition footage. As a result, video is likely the most cost-effective way of creating multimedia that gives the learner control over the pace of their own learning. While computer games and simulations might also be multimedia that offer students control, they are significantly more resource intensive than offering students a video with the capacity to control the playback. From the perspective of cognitive load theory, this control allows students to regulate the intrinsic load of the task—that is, the load on their working memory inherent in the task itself (Van Gog et al., 2005; van Merriënboer & Sweller, 2005). For example, calculating a t value requires that the student hold the formula in their head alongside the means and pooled standard deviation. User control also allows for better management of germane cognitive load—where students invest resources constructing schema or developing automaticity—by, for example, taking notes or developing concept maps as they go (Chi & Wylie, 2014; Van Gog et al., 2005; van Merriënboer & Sweller, 2005). To use the previous example, the video could be paused so the student could connect a t value to a correlation coefficient. Face-to-face classes and videoconferences could also be paused for these activities, but videos allow students to control the amount of time they invest.

By being asynchronous, videos allow the teacher to have more control over the presentation via editing. Any multimedia—including face-to-face classes and videoconferences—can be crafted such that they reduce extraneous load by applying a series of multimedia design principles (Mayer, 2009; Mayer et al., 2020; Mayer & Moreno, 2003). Extraneous load is where working memory and attention are “wasted” on content that is not inherent to the learning activity (van Merriënboer & Sweller, 2005). Again, most academics can relate to listening to a lecture that digresses along winding, tangentially relevant monologues. These “seductive details” may be interesting to some students, but meta-analyses have shown that irrelevant information reduces learning because these details use working memory that would otherwise be better focused on the core learning activity (coherence principle; Mayer, 2008; Rey, 2012). These details are easier to limit when teaching staff have the ability to edit videos before dissemination. Similarly, lab studies have shown multimedia to be more effective when important details are highlighted (signaling effect; Mayer, 2008; S. Schneider, Beege, et al., 2018), when words and pictures are presented at the same time (temporal contiguity; Ginns, 2006; Mayer, 2008), and as part of the same visual field (spatial contiguity; Ginns, 2006; Mayer, 2008). All of these multimedia design principles could—theoretically—also be implemented in face-to-face classes and synchronous videoconferences (e.g., via spoken lecture paired with visual slide deck). In the same way, a writer could theoretically produce a flawless piece of prose on their first pass, but the first draft of any piece usually benefits from editing. Implementing these instructional methods is easier when teaching staff have the ability to edit and review videos before dissemination.

While learner pacing and editing may be arguments in favor of video, interactivity may be an argument against it. It is well established that student learning is proportional to the amount of interactivity (Bernard et al., 2009; Chi & Wylie, 2014). Passive viewing has been shown to be less effective than active engagement (e.g., taking notes), which is less effective than constructive processing (e.g., generating a concept map), which is less effective than co-construction with another learner (e.g., co-construction of the concept map; Chi et al., 2018; Chi & Wylie, 2014). Online, video does not necessarily prohibit interactivity, but interactivity is facilitated in an environment where students are working together in the same classroom or virtual break-out room. As a result, video may be less effective than synchronous classes (face-to-face and videoconference) if all three effectively manage cognitive load, but the synchronous teaching methods promote more interactivity.

This review aimed to assess the effects of video on learning in higher education. Assuming the interactivity was held constant, we hypothesized that students given video would be as good as other forms of instruction and possibly beneficial due to editing and student self-pacing. If video was presented in a less interactive context, we hypothesized that it would be less effective. Due to the aforementioned multimedia effect, we hypothesized that videos would be superior to static, asynchronous media, like textbooks or static websites. We conducted a systematic review including studies that compared video with any comparison condition. We included studies that swapped videos for their existing content and those that added videos in addition to existing instruction. We analyzed these two types of studies separately to avoid pooling heterogeneous comparison conditions. While it is reasonable to hypothesize that more content is likely to be better, we sought to assess the size of these effects. This is because there may be many opportunities where universities have the ability to provide supplemental video content, and these studies would allow staff to determine whether or not more content is necessarily better. We included only randomized trials because they usually allow for stronger causal inference. Quasi-experimental studies are more likely to be confounded, particularly if students could choose the mode of delivery that suited them best. We reviewed the effects on learning, rather than student satisfaction, because learning is a key goal of all universities, and student evaluations are poor predictors of educational attainment (Uttl et al., 2017, 2019). Overall, in this review we aimed to determine (a) what are the learning effects of swapping other content for videos and (b) what are the effects of adding videos to existing course materials?

Method

We prospectively registered this systematic review via PROSPERO (CRD42016046173, see Supplementary File 1, available in the online version of the journal). Presentation in this manuscript is aligned with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA statement; Moher et al., 2010) and the Reporting Standards for Research in Psychology (APA Publications and Communications Board Working Group on Journal Article Reporting Standards, 2008). These reporting standards help ensure that the conclusions of systematic reviews are robust and reproducible by reporting key decisions made in the review (Moher et al., 2010). The standards also guide reviewers toward critical components that are often neglected in systematic reviews, such as publication bias (Moher et al., 2007). Educational research often necessitates different expectations and priorities, so we also adhered to recommended processes for high-quality systematic reviews of educational research (Alexander, 2020; García-Holgado et al., 2020; Pigott & Polanin, 2020).

Eligibility Criteria

We selected studies on the basis of prespecified inclusion criteria:

1. *Design:* We included randomized designs because of their stronger ability to make causal inferences; specifically, studies needed some form of randomization at either an individual or a cluster level, or randomized counterbalancing in the case of repeated measures designs. As a result, we excluded nonrandomized trials or cohort studies.
2. *Interventions:* We included interventions where the independent variable was the use of video for teaching and learning (e.g., lecture capture, educational multimedia). Studies were still eligible if students also learned content via other methods (e.g., tutorials) as long as those methods were consistent across groups (e.g., both groups given the same tutorials, but lectures were either face-to-face or video). Video did not need to be the only form of instruction, but video did need to be the primary difference between experimental arms. We therefore excluded studies where video could not be isolated as the independent variable, such as when adding video was confounded with other significant changes (e.g., flipped classrooms where lectures are moved to video but students are also given a substantially different learning experience via interactive workshops). Synchronous multimedia with live interaction from an instructor (e.g., videoconferencing) was excluded because it is qualitatively different from prerecorded, asynchronous video, with different costs and benefits, as outlined earlier.

3. *Comparisons:* We included comparisons against any other types of teaching and learning (e.g., assigned readings, face-to-face lectures), as well as the addition of video to supplement existing educational methods (e.g., lectures vs. lectures + recordings). We analyzed these two scenarios separately in moderation analyses.
4. *Outcome:* Studies needed to report a measure of learning or achievement (usually a measure of knowledge or skills); we did not include studies that only include student satisfaction, tolerability, or self-report measures of perceived learning or achievement.
5. *Participants:* To maintain a homogeneous population, studies needed to gather data from students enrolled in a higher education institution (i.e., undergraduate or postgraduate students at university or college); we did not include studies that only reported primary students, secondary students, professional learners, or a combination of the above.
6. Finally, studies needed to be in English or have an available English translation, and needed to present original data (i.e., we excluded review articles). Published and unpublished studies from any time period were eligible for inclusion.

Information Sources and Search Strategy

To ensure the search strategy was comprehensive (Alexander, 2020), we generated a search strategy using the titles and abstracts of an initial sample of papers, including existing reviews (Adesope & Nesbit, 2012; Ginns, 2006; Höffler & Leutner, 2007; Kay, 2012; Mayer & Moreno, 2003; Rey, 2012; Shrivastav & Hiltz, 2013) and primary studies (Baxter et al., 2012; Muller et al., 2007; Muller et al., 2008; Rath & Holt, 2010; Steedman et al., 2012). We generated a list of terms that identified the target papers (Eriksen & Frandsen, 2018) optimizing for sensitivity while maintaining specificity. The list of search terms was as follows:

- *Participants*: (university *or* undergraduate *or* "tertiary education" *or* "higher education" *or* "student") *and*
- *Intervention*: (video *or* "video learning" *or* "e-learning" *or* "educational video" *or* "instruct* video" *or* "online lectures" *or* multimedia *or* "lecture-style presentation") *and*
- *Comparison*: (randomised *or* randomized *or* randomly *or* trial *or* groups) *and*
- *Outcome*: (results *or* test *or* grades *or* marks *or* "examination results" *or* assessment *or* "academic performance" *or* "academic achievement" *or* "educational outcomes")

This set of terms was entered into five databases: ERIC, PsycINFO, EMBASE, MEDLINE, and CINAHL. We conducted searches on August 23, 2019.

Study Selection

First, we removed duplicates in EndNote (The EndNote Team, 2013). Then, to screen against inclusion criteria, one of two reviewers screened titles and abstracts as “possibly relevant” or “irrelevant.” After screening, the lead author audited the “irrelevant” articles for any mis-classifications using DistillerAI (https://www.evidencepartners.com/products/distillersr-systematic-review-software/). DistillerAI used both a Naive Bayes classifier and a Support Vector Machine classifier to score references in terms of their likelihood for inclusion. Excluded references scoring above the default threshold (probability of relevance = 0.5) were flagged for rescreening and moved to full-text screening if possibly relevant. Each full-text article was then independently assessed against the inclusion/exclusion criteria in duplicate, with conflicts resolved through discussion or consultation with a third reviewer. Reviewers logged reasons for excluding any potentially relevant studies. One reviewer searched the reference lists of included studies for papers that may have been missed by the search strategy (Pigott & Polanin, 2020). Studies were included in the meta-analysis when we were able to either extract or impute sufficient data for effect size calculations, following guidelines from the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins et al., 2019).

Data Items and Collection Process

Two reviewers developed, piloted, and revised a data extraction form. The form extracted details regarding the participants (subject students were studying; region), learning context (lectures, tutorials, homework, or a combination), interventions (description, duration, level of active learning, type of video [e.g., skills demonstration, recorded lecture]), comparison (description, duration, level of active learning, whether teacher or not), outcome (description, whether knowledge or skills, follow-up interval), and funding information (whether or not funding was reported, source of funding). We also extracted any metric that could be used to calculate an effect size (e.g., means, standard deviations, confidence intervals, *p* values). We independently extracted all items in duplicate and resolved disagreements via discussion, with adjudication by a third author, when required.

Risk of Bias in Individual Studies

We assessed risk of bias in individual studies using the Cochrane Risk of Bias Assessment for randomized trials (Higgins et al., 2011). This tool assesses whether studies implemented strategies to prevent six categories of bias: selection bias, performance bias, detection bias, attrition bias, reporting bias, and other biases. While we acknowledge this is a stringent standard to apply, we chose it for a number of reasons. The Cochrane tool has demonstrated greater sensitivity, specificity, and validity compared with other quality assessment tools (Higgins et al., 2011). Failing to meet each criterion has been shown to bias results in meta-meta-analyses (Higgins et al., 2011). This contrasts with many quality checklists that conflate issues of reporting (e.g., reporting the inclusion criteria) with issues of bias (e.g., unblinding the outcome assessors). To create an overall rating for each study, the Cochrane tool does not recommend adding the number of criteria met by each study. While calculating a “quality score” offers readers parsimony, doing so misleadingly suggests that each criterion is equally and incrementally problematic (Higgins et al., 2011). Instead, reviewers used the criteria for "overall risk of bias" from the *Cochrane Handbook* (Higgins et al., 2011) in which studies are considered low risk if they are judged to be low risk on each domain. We conducted all quality assessments independently and in duplicate, with disagreements resolved through discussion and consultation with a third author, where necessary.

Summary Measures and Synthesis of Results

We extracted data for each effect size reported within each study, because merely averaging multiple measures leads to biased estimates of variance (Moeyaert et al., 2017). The principal summary measure was the posttest, between-groups standardized mean difference. We used Hedges's g, which corrects for biases in small sample sizes (Hedges, 1981). This measure was calculated using the *metafor* package (Viechtbauer, 2017) in *R* (R Core Team, 2020). Interpretation for Hedges's g is the same as for Cohen's d: Many authors use Cohen's rule-of-thumb (small = 0.2; moderate = 0.5; large = 0.8) or effects could be compared with other studies in the field (Durlak, 2009). For example, the median effect size for meta-analyses in higher education is .35 (M. Schneider & Preckel, 2017).

Omitting trials due to missing data on some variables is generally not recommended (Higgins et al., 2019). Where means and standard deviations were not available, we used other statistics to calculate effect sizes (e.g., posttest confidence intervals, p values) or imputed data using recommendations and formulae from the *Cochrane Handbook* (Higgins et al., 2019).

We conducted multilevel meta-analyses (Moeyaert et al., 2017), nesting effect sizes within studies, using the *metasem* package (Cheung, 2014) and *msemtools* (Conigrave, 2019) in *R* (R Core Team, 2020). For each analysis, we assessed heterogeneity using I^2 at Level 2 and Level 3 (within and between studies, respectively), which indicates the percentage of variance that is unlikely to be merely an artifact of sampling error (Higgins et al., 2011). We also assessed heterogeneity using processes outlined by Mathur and VanderWeele (2019a), which looked at the proportion of true effects likely to be helpful (which we defined as a small, positive effect; Hedges's $g > 0.2$) or harmful (Hedges's $g < -0.2$).

Additional Analyses

We conducted a series of moderation analyses to explore possible sources of heterogeneity. These included (a) educational setting (i.e., lectures, tutorials, homework); (b) comparison condition (teacher vs. static media); (c) type of outcome assessment (knowledge test vs. skills assessment); (d) outcome timing (immediately post intervention vs. after some follow-up interval); (e) the relative length of the interventions (i.e., whether video was matched to control, longer than control, or shorter than control); (f) the relative interactivity of the intervention (i.e., whether video was matched to control, more active than control, or less active); and (g) the absolute length of the intervention, both via meta-regression using minutes to predict effect size, and by whether the intervention was a single topic or a full course. We conducted sensitivity analyses to assess differences between studies that met each criterion on the Cochrane Risk of Bias tool. For these sensitivity analyses, we grouped “unclear” and “high” risk studies together, as recommended by the *Cochrane Handbook*, due to their similar profiles of risk (Higgins et al., 2011).

Risk of Bias Across Studies

To assess publication bias, we inspected a funnel plot and conducted a three-parameter selection model (3PSM; Hedges & Vevea, 1996). We chose this method because simulation studies have shown 3PSM demonstrates better management of Type I and Type II errors compared with other assessments of publication bias (e.g., Egger’s regression test, rank correlation test, Trim-and-Fill; Pustejovsky & Rodgers, 2019). The 3PSM identifies whether studies are more likely to be published when significance values are within certain categories (e.g., $p < .05$ vs. $p > .05$). A significant likelihood ratio for this test indicates the presence of publication bias, because it means that effect sizes in some of these brackets are more likely to be published than others. In addition, we conducted sensitivity analyses for publication bias described by Mathur and VanderWeele (2019b). These analyses estimate how strong publication bias would need to be to nullify pooled effect size estimates. In other words, how much more likely must significant results be to account for the observed effects?

Results

Study Characteristics

The full results of the study selection process are outlined in the PRISMA flow diagram (see Figure 1). After database screening and removal of duplicates, we found 9,670 studies. We located an additional seven studies by searching the reference lists of included studies. After screening, 325 studies were marked as possibly relevant, and an additional four (0.04%) were identified as possibly relevant following the DistillerAI audit. We assessed 329 full-text articles in duplicate against eligibility criteria. Of those, we excluded most (134) because they were not randomized trials (e.g., retrospective studies; Farooq & Al-Jandan, 2015). We excluded 41 because they did not have video or multimedia as their independent variable (e.g., reflective e-journals only; Chang & Lin, 2014). Other reasons for exclusion included the following: 22 because students were not in college or university (e.g., high school students; Beydogan & Hayran, 2015); nine had confounded designs with more than video as an independent variable (e.g., flipped classrooms; Albalawi, 2018); eight did not report an achievement outcome (e.g., feelings of stigma only; Fernandez et al., 2016); eight were reviews (e.g., Tularam & Machisella, 2018); and two were duplicates of included studies (O’Donovan et al., 2016; Sayed & Abdelmonem, 2018). This left a combined total of 105 studies and a pooled sample size of 7,776 students. The studies were overwhelmingly conducted in Western countries, with most in North America (44%), Europe (25%), or Oceania (5%). The rest were spread across the Middle East (12%), Asia (10%), South America (3%), and Africa (1%). The specifics of each study are listed in the Characteristics of Included Studies table (Supplementary File 3, available in the online version of the journal).

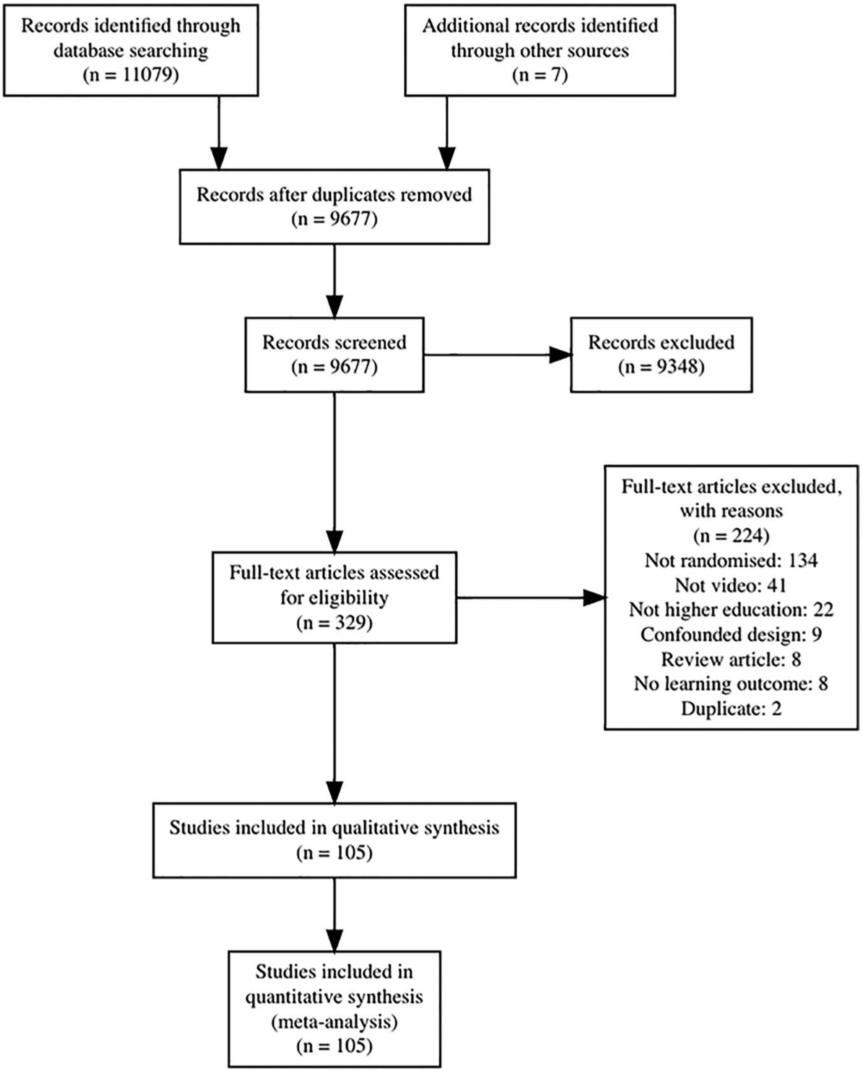

FIGURE 1. *PRISMA flow diagram for inclusion in this review.*
Note. PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-analyses.

Effects of Replacing Other Teaching With Video

The effect sizes extracted from individual studies are available in the forest plot in Figure 2. Full characteristics of each included study are presented on the Open Science Framework for transparency and reproducibility of analyses (bit.ly/betteronyoutube). Overall, replacing other teaching with video had a significant positive effect on student learning ($g = 0.28$, 95% CI [0.14, 0.42], $n = 166$,

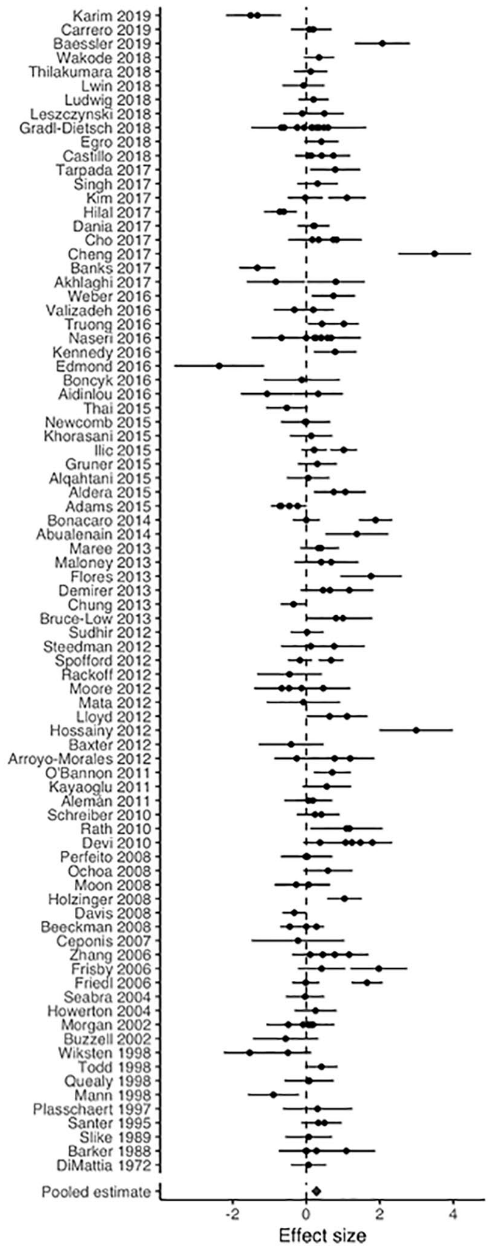

FIGURE 2. *Effect size for each study that compared swapping video for any other learning opportunity.*

Note. Each point reflects an effect size (with confidence interval) within each study. Studies where video was added to existing content are presented later. Effect sizes are ordered according to date of publication, from the newest to the oldest study.

$k = 83$). This point estimate demonstrated a high level of heterogeneity that was not explained by sampling error alone ($I_2^2 = 0.32$, $I_3^2 = 0.55$, $\tau_2^2 = 0.16$, $\tau_3^2 = 0$). As mentioned earlier, we explored this heterogeneity by assessing the likely distribution of true effect sizes; we identified the proportion of true effects that would be helpful (i.e., $g > 0.2$) or harmful (i.e., $g < -0.2$). Based on effects found here, half the implementations of exchanging other learning for video will be helpful (50% of true effects, 95% CI [40%, 59%]). A small proportion of implementations may be unhelpful for student learning (19% of true effects, 95% CI [13%, 25%]) with the rest having negligible influences.

Moderation Analyses by Participants, Setting, and Comparison Condition

The results of all moderation analyses are presented in Table 1. When examining the effects of video interventions across different context, region ($R_2^2 = 0.00$, $R_3^2 = 0.02$, $p = .98$) was not a significant moderator of effects. Effects were not moderated by the subject being studied, regardless of whether conceptualized broadly (e.g., health sciences vs. language learning; $R_2^2 = 0.00$, $R_3^2 = 0.02$, $p = .85$) or narrowly (e.g., nursing vs. medicine vs. dentistry; $R_2^2 = 0.00$, $R_3^2 = 0.25$, $p = .79$). Funding information was not a significant moderator, either. Effects were not moderated by the kind of video being used (i.e., recorded lectures vs. case examples; $R_2^2 = 0.01$, $R_3^2 = 0.01$, $p = .95$). Effects were not significantly different between studies that used videos in lectures, tutorials, or homework ($R_2^2 = 0.00$, $R_3^2 = 0.18$, $p = .07$). In contrast, the comparison condition was a significant moderator ($R_2^2 = 0.00$, $R_3^2 = 0.17$, $p = .03$): When video replaced static media (e.g., text), it was significantly more effective ($g = 0.51$, 95% CI [0.27, 0.75], $n = 45$, $k = 27$) than when video replaced a teacher ($g = 0.18$, 95% CI [0.02, 0.34], $n = 121$, $k = 58$).

Moderation Analyses by Outcome Assessment

The type of outcome measured was a significant moderator ($R_2^2 = 0.15$, $R_3^2 = 0.00$, $p = .02$). Video was more effective when students were assessed on skill acquisition ($g = 0.44$, 95% CI [0.24, 0.64], $n = 67$, $k = 38$) compared with assessments of their knowledge ($g = 0.18$, 95% CI [0.01, 0.35], $n = 99$, $k = 60$). We also assessed for moderation by whether the assessment was directly after the intervention, or after a follow-up period. The timing of the assessment did not significantly moderate intervention effects ($R_2^2 = 0.00$, $R_3^2 = 0.00$, $p = .85$).

Moderation Analyses by Interactivity and Duration of Content

The relative amount of educational content did not significantly moderate the benefits of videos for learning ($R_2^2 = 0.01$, $R_3^2 = 0.07$, $p = .18$). Students learned roughly the same when the video condition received more training than control, when control received more than video, and when conditions were matched. The absolute amount of content was not a significant moderator either. The number of minutes of the educational intervention did not moderate effects ($b = 0.00$, 95% CI [0.00, 0.00], $p = .73$), and there were no significant differences in effects when the video intervention was applied to a single topic or a whole course ($R_2^2 = 0.00$, $R_3^2 = 0.00$, $p = .08$). In other words, there was no significant dose–response effect.

TABLE 1

Moderation analyses for swapping other educational content for videos

Moderator	k	n	Estimate [95% CI]	SE	$R^2_{(2)}$	$R^2_{(3)}$	p
Baseline ($I^2_{(2;3)}$: 0.32; 0.55)	83	166	0.28 [0.14, 0.42]	0.07	—	—	—
Educational setting	83	166	—	—	0.00	0.18	.068
Comparison	83	166	—	—	0.00	0.17	.027*
Static media	27	45	0.51 [0.27, 0.75]	0.12	—	—	—
Teacher	58	121	0.18 [0.02, 0.34]	0.08	—	—	—
Outcome	83	166	—	—	0.15	0.00	.019*
Knowledge test	60	99	0.18 [0.01, 0.35]	0.08	—	—	—
Skills assessment	38	67	0.44 [0.24, 0.64]	0.10	—	—	—
Outcome timing	83	166	—	—	0.00	0.00	.850
Which is longer	83	166	—	—	0.01	0.07	.184
Which is more interactive	83	166	—	—	0.01	0.12	.038*
Control was more interactive	16	32	–0.07 [−0.38, 0.25]	0.16	—	—	—
Interactivity was equivalent	60	126	0.33 [0.17, 0.49]	0.08	—	—	—
Video was more interactive	8	8	0.62 [0.13, 1.11]	0.25	—	—	—
Topic or course	82	161	—	—	0.00	0.00	.076
Kind of video	83	166	—	—	0.01	0.01	.95
Broad learning domains	83	166	—	—	0.00	0.02	.85
Narrow learning domains	83	166	—	—	0.00	0.25	.79
Region	83	166	—	—	0.00	0.02	.98
Funding reported	83	166	—	—	0.00	0.01	.58
Funding source	83	166	—	—	0.00	0.13	.44

Note. k = number of studies; n = number of effects from those studies; Educational setting = whether video replaced content in lectures, tutorials/labs, homework, or a mixture. Outcome type = whether outcome assessment focused on knowledge or skill acquisition. Outcome timing = whether outcome assessment was immediately after the intervention or later (e.g., end of semester). Topic or course = whether the intervention spanned one topic or an entire course).
$^*p < .05$.

The relative interactivity was a significant moderator of effects ($R_2^2 = 0.01$, $R_3^2 = 0.12$, $p = .04$). There was no benefit to video when the control condition was afforded more interactivity ($g = -0.07$, 95% CI [−0.38, 0.25], $n = 32$, $k = 16$). Videos were effective when both conditions were given equivalent opportunities for interactivity ($g = 0.33$, 95% CI [0.17, 0.49], $n = 126$, $k = 60$). Effects were particularly large when videos were presented in an interactive context (e.g., coviewing with a peer) that was not available to the control condition ($g = 0.62$, 95% CI [0.13, 1.11], $n = 8$, $k = 8$).

Effects of Providing Supplementary Videos

Figure 3 displays effect sizes for interventions that provided videos in addition to existing content. There was a strong, significant effect of providing students with supplemental videos ($g = 0.88$, 95% CI [0.62, 1.13], $n = 64$, $k = 34$). This point estimate demonstrated a high level of heterogeneity that was unexplained by sampling error alone ($I_2^2 = 0.55$, $I_3^2 = 0.37$). Regardless, almost all (88% of true effects, 95% CI [75%, 94%]) implementations of adding video are expected to be helpful for student learning ($g > 0.2$). Effectively no implementations of supplementary videos would be bad for learning (2% of true effects, 95% CI [0%, 5%]).

As shown in Table 2, there were no differences between the effects of adding video for teaching knowledge or skills ($R_2^2 = 0.00$, $R_3^2 = 0.28$, $p = .22$). The effects were not moderated by whether the assessment was directly after the content or after some follow-up period ($R_2^2 = 0.01$, $R_3^2 = 0.06$, $p = .32$). The effects were not significantly different when videos were added to one topic or throughout the whole course ($R_2^2 = 0.02$, $R_3^2 = 0.02$, $p = .34$). The combined length of the videos condition (in minutes) did not moderate the effects ($b =$ 0.00, 95% CI [0.00, 0.00], $p = .53$). The region in which the study was conducted was a significant moderator ($R_2^2 = 0.08$, $R_3^2 = 1.00$, $p < .001$), but the small number of studies in most regions (see Table 2) may mean these models are overfitting. Specifically, only two regions had more than three studies—Europe ($g = 0.82$, 95% CI [.48, 1.15], $n = 7$, $k = 14$) and North America ($g = 0.71$, 95% CI [.47, .95], $n = 18$, $k = 27$)—and pooled effects from these two regions did not differ. Funding was a significant moderator of effects ($R_2^2 = 0.00$, $R_3^2 = 0.37$, $p = .011$). Projects that received funding demonstrated significantly higher effect sizes ($g = 1.22$, 95% CI [0.88, 1.56], $n = 14$, $k = 36$) than those that were unfunded ($g = 0.59$, 95% CI [0.28, 0.91], $n = 20$, $k = 28$).

In all studies examining the effect of additional video, the two conditions were matched except one condition was given additional video resources (content vs. content + videos). As a result, all other moderators from the previous analyses were less meaningful. For example, the video condition was always longer, so the relative length of the intervention was unnecessary. We therefore did not conduct any further moderator analyses.

Sensitivity Analyses Removing Imputed Standard Deviations

As mentioned earlier, we calculated effect sizes using all available data. Where no useful information was available to calculate the variance, we imputed standard deviations using a conservative estimate (90th percentile, as per Higgins et al., 2019) because this generally leads to less biased estimates compared with

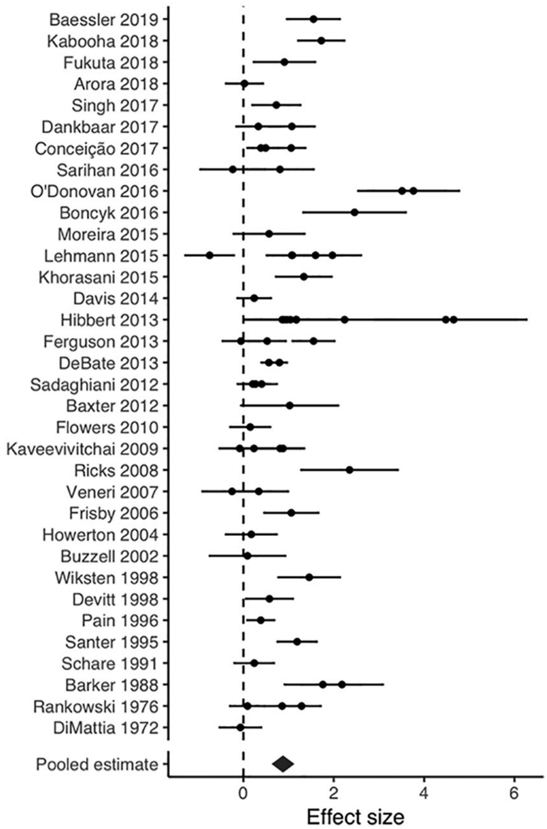

FIGURE 3. *Effect size for each study that provided videos in addition to existing content.*
Note. Each point reflects an effect size (with confidence interval) within each study. Effect sizes are ordered according to date of publication, from the newest to the oldest study.

removing the studies. We conducted a sensitivity analysis for these imputed standard deviations by running models omitting these effect sizes and instead using full information maximum likelihood (Enders, 2010). In these analyses, the pooled estimates were basically unchanged for swapping content for both video (g = 0.28, 95% CI [0.13, 0.43], n = 154, k = 78) and adding additional videos (g = 0.84, 95% CI [0.59, 1.09], n = 59, k = 32). All moderation analyses followed the same pattern, with one exception. When swapping content for video and omitting imputed standard deviations, the comparison condition was no longer a significant moderator ($R_2^2 = 0.00$, $R_3^2 = 0.15$, $p = .06$).

Risk of Bias Within Studies

The consensus ratings of risk of bias for each study are presented in Supplementary File 3 (available in the online version of the journal). Overall, no

TABLE 2

Moderation analyses for adding videos to other educational content

Moderator	k	n	Estimate [95% CI]	SE	$R^2_{(2)}$	$R^2_{(3)}$	p
Baseline ($I^2_{(2;3)}$: 0.55; 0.37)	34	64	0.88 [0.62, 1.13]	0.13	—	—	—
Outcome type	34	64	—	—	0.00	0.28	.216
Outcome timing	34	64	—	—	0.01	0.06	.316
Topic or course	34	64	—	—	0.02	0.02	.345
Kind of video	34	64	—	—	0.00	0.08	.93
Broad learning domains	34	64	—	—	0.06	0.56	.38
Narrow learning domains	34	64	—	—	0.03	0.18	.36
Region	34	64	—	—	0.08	1.00	<.001*
Oceania	2	11	1.45 [1.12, 1.77]	0.17	—	—	—
Europe	7	14	0.82 [0.48, 1.15]	0.17	—	—	—
North America	18	27	0.71 [0.47, 0.95]	0.12	—	—	—
Asia	3	6	0.43 [−0.08, 0.93]	0.26	—	—	—
Middle East	3	4	0.95 [0.29, 1.60]	0.33	—	—	—
Africa	1	2	3.63 [2.56, 4.71]	0.55	—	—	—
Funding reported	34	64	—	—	0.00	0.37	.011*
Funding reported	14	36	1.22 [0.88, 1.56]	0.18	—	—	—
No funding reported	20	28	0.59 [0.28, 0.91]	0.16	—	—	—
Funding source	34	64	—	—	0.10	0.90	<.001*
Federal government grant	3	6	0.58 [0.05, 1.12]	0.27	—	—	—
Internal university grant	6	23	1.14 [0.80, 1.47]	0.17	—	—	—
No funding described	20	28	0.58 [0.33, 0.83]	0.13	—	—	—
Nonprofit foundation	2	3	2.76 [1.89, 3.64]	0.44	—	—	—
Private industry grant	1	1	0.58 [−0.72, 1.87]	0.66	—	—	—
State government grant	2	3	1.68 [0.79, 2.56]	0.45	—	—	—

Note. k = number of studies; n = number of effects from those studies. Outcome type = whether outcome assessment focused on knowledge or skill acquisition. Outcome timing = whether outcome assessment was immediately after the intervention or later (e.g., end of semester). Topic or course = whether the intervention spanned one topic or an entire course.
*$p < .05$.

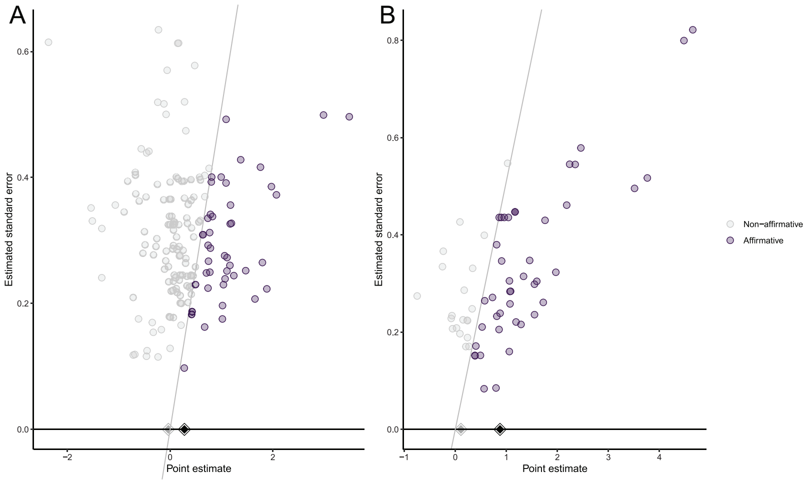

FIGURE 4. *Funnel plot visualizing effect sizes proportional to standard errors.*
Note. Panel A shows the effects of swapping content for videos, and Panel B shows the effects of adding video to content. Dark points to the right of the line indicate affirmative studies (i.e., those showing significant benefits of video); light points to the left of the line indicate nonaffirmative studies (i.e., nonsignificant results or significant negative effects). Diamonds at the bottom indicate the pooled effect size estimates using all studies (black) and the worst-case (light grey), using only nonaffirmative studies. Results are still positive and significant for adding video, but benefits of swapping content for video are reduced to near-zero when only including non-affirmative studies.

studies were evaluated as low risk of bias; all were either unclear (k = 62) or high risk (k = 43). This was largely because most studies failed to prospectively register their outcomes, with only one study being evaluated as low risk for selective reporting (Ilic et al., 2015). For selection bias, 36 were low risk for sequence generation and 22 for allocation concealment. For performance bias, only eight blinded participants and 31 blinded personnel. Most studies (k = 82) used a standardized assessment (e.g., multiple-choice questionnaire or blinded evaluation of skills) such that outcome measurement was unlikely to be biased. Most studies (k = 72) either had minimal missing data or adequately handled data with intention-to-treat principles. Only five studies demonstrated other risks (e.g., systematic differences in assessment conditions). In sensitivity analyses, none of the risk of bias criteria explained significant variance in effect sizes (see Supplementary Tables 1 and 2, available in the online version of the journal).

Risk of Bias Across Studies

We plotted effect sizes against standard errors using the funnel plots (Figure 4). The 3PSM likelihood ratio test indicated a pattern of effect sizes consistent with publication bias ($\chi^2(7) = 35.50, p < .001$), such that affirmative studies appeared more likely to be published. However, a strong publication bias would be required to nullify these effects. For the effects of swapping content with video to be drawn to zero, significant studies would have to be 10.48 times as likely to be published (Lower bound of CI: 2.41 times as likely). For supplemental video, no level of publication bias could reasonably account for the significant effects. As seen from Figure 4b, worst-case analyses, including only nonsignificant studies, still leads to a positive pooled effect.

Discussion

As universities move toward online learning through videos, some academics may fear that students will perform less well compared with their peers who receive traditional methods. Other things being equal, our review shows student results are unlikely to decline when using video for teaching. We identified more than 100 randomized trials that had used video in higher education, and on average, videos led to better learning outcomes compared with other methods. In the 83 studies that swapped existing learning for videos, there were small learning benefits, with meaningful positive effects anticipated in about half the implementations of video. These results are consistent with previous studies of online learning (broadly defined) that found it to be as good, and sometimes better, than face-to-face teaching (Means et al., 2009, 2010; Means et al., 2013). Our results were robust across different settings (e.g., lectures or tutorials), different domains (e.g., science or languages), different types of video (e.g., case demonstrations or recorded lectures), different length interventions, and different follow-up periods. Most of these results, and the moderation analyses, can be explained by the cognitive theory of multimedia learning (Mayer, 2008), cognitive load theory (van Merriënboer & Sweller, 2005), and the ICAP (interactive, constructive, active, passive) framework (Chi & Wylie, 2014).

Findings May Be Explained by the Cognitive Theory of Multimedia Learning

Multimedia learning can make optimal use of cognitive architecture by providing coherent learning stimuli through visual and aural channels (Mayer, 2008; Rolfe & Gray, 2011). Many forms of instruction—including video, face-to-face classes, and videoconference—can be optimized for this architecture, but some forms of instruction (e.g., textbooks, podcasts) generally only operate via one channel. Our moderation analyses are consistent with this interpretation, where video is superior to nonteacher comparison conditions (e.g., textbooks) and more similar to university teachers.

The finding that video was superior to even face-to-face classes may be explained in a few ways. It may be due to the increased ability for students to manage their own cognitive load (e.g., by pausing and rewinding) or because teachers can better optimize cognitive load through editing. We could not test these hypotheses because few studies reported whether the video was self-paced or other-paced, or whether the video was edited. We suspect that both effects play a part, for reasons described below.

When learners cannot control the pacing, multimedia design principles become more important, because students can no longer compensate for bad design by taking control. For example, Adesope and Nesbit (2012) found that putting key points on slides or videos ("redundancy principle") was *more* important when the presentation was other-paced. This is because, as the learner falls behind, they can look at those key points to identify gaps. Without them the learner cannot catch up if the video is other-paced. When self-paced, the learner could rewind to review material that was otherwise lost, and the benefit of the redundancy effect was eliminated (Adesope & Nesbit, 2012). Similarly, the benefits of hearing content alongside graphics—as opposed to reading text alongside graphics ("modality principle"; Mayer, 2008)—was significantly reduced when the media was self-paced (Ginns, 2005). Again, when students can rewind content, they can review at their own pace, even though hearing the content would be more efficient. With these findings in mind, it may be reasonable to assume that many of the implementations of video in our review allowed for the student to have some control over the video's playback and that this control allowed students to regulate their cognitive load, leading to better learning. This ability to self-pace has been previously identified as a key feature contributing to the success of online learning, more generally (Tallent-Runnels et al., 2006).

Another mechanism by which video may be advantageous is due to the teacher's ability to edit. With editing, they can make content more coherent, and add design principles that they would not perfectly execute in class (e.g., timing key points with slides; highlighting important information). Our results suggest that video was more effective than comparison conditions, even when the comparison conditions were longer. We find video is more time efficient, with the same content condensed into a shorter period (e.g., a 2-hour lecture condensed into a 40-minute video). Assuming no differences in interactivity (addressed later), concise videos may teach more effectively than longer classes because of the coherence principle (a.k.a., "seductive detail effect," described earlier; Rey, 2012). That is, these videos may force educators to prioritize core content, editing out tangential details that are not important for the learning objectives. Doing so may explain why videos often perform better. All of these principles would apply regardless of the domain being studied, the context of the video, or the kinds of video being presented.

Finally, video may have stronger effect sizes than other teaching methods because it is able to provide a different, more authentic perspective. This may explain why videos were more effective for teaching skills than transmitting knowledge. When learning the history of feminism in India, for example, there may be no substantive difference between sitting in a class and watching a video. But when learning about cardiopulmonary resuscitation, a medical procedure (e.g., heart surgery), or a counselling skill (e.g., suicide assessment), students will (hopefully) only see contrived procedures when taught in class. They may also see the model procedure from a distance because they are in a room with dozens of other students. In contrast, videos allow for students to see authentic demonstrations of skills with real people. They also allow for unique perspectives where students can see a skill through the eyes of the performer. It may be a similar mechanism by which virtual reality experiences demonstrate promising learning outcomes (Kyaw et al., 2019; Radianti et al., 2020). Virtual reality is much more expensive and difficult to implement, but it affords students a higher level of interactivity compared with typical videos. While comparisons of effect sizes between meta-analyses are fraught (M. Schneider & Preckel, 2017), this interactivity may explain the large benefits of virtual reality compared with other types of instruction (Kyaw et al., 2019).

Videos were more effective when provided in an environment that was as interactive—or more interactive—than the traditional teaching methods. In studies where traditional methods were more interactive, video was no longer beneficial. These results are consistent with previous studies on interaction in distance learning (Bernard et al., 2009), and the ICAP framework that suggested learning is proportional to interactivity (Chi & Wylie, 2014). Interaction can take many forms, with learning benefits when students interact with teachers, content, or each other (Bernard et al., 2009). While some types of interactivity are easier in classes or videoconferences (e.g., student–student peer discussion; student–teacher question-and-answer), many face-to-face classes are passive and many video-based interventions are constructive (e.g., Khan Academy demonstrates rich student–content interaction with questions and feedback). Video is typically "passive" or "active" (with only an ability to take notes, pause, and rewind), but so are traditional lectures (with the ability to take notes, etc.). Just as discussions and activities can increase the interactivity in lectures, online discussions and embedded questions (e.g., using H5p or EdPuzzle) can increase opportunities for constructive learning in multimedia. Our results suggest that moving content to video is more effective when teachers maintain, or even increase, the level of active learning offered to students.

Our review did not find a dose–response effect. That is, there was no systematic difference between the learning benefits when videos were applied in brief experiments or across a whole semester. Dose–response effects are difficult to detect in most meta-analyses and are best assessed when participants are randomized to one dose or another (Deeks et al., 2011). Random assignment to different dosages is rare in educational research, and even rarer when assignment must be done at the unit or university level (rather than at the student level). Few of the studies we included involved two arms with different doses, so our abilities to make strong conclusions here are limited. Having said that, the absence of a dose–response may be due to the ways students supplement their learning during longer studies. Over time, the benefits of video may be counterbalanced by other types of learning, including self-directed revision or constructively aligned assessments (Biggs & Tang, 2011).

Overall, video is ultimately a medium and the medium itself is indeed the only space in which students learn (Clark, 1994; Warnick & Burbules, 2007). However, video is a medium that, in many ways, enables teachers to more easily use important instructional methods (Kozma, 1994; Warnick & Burbules, 2007). Editing videos allows teachers to more easily implement a host of multimedia design principles, each of which are beneficial for learning. Instructional technologies are now enabling videos to be as interactive, if not more interactive, as many traditional forms of teaching. Video allows students to engage in the content (and the interactivity) at their own pace and in their own time. It is true that video is unlikely to be optimal on its own, and it is probably more effective when used in combination with dynamic student–student and student–teacher interactions. However, shifting didactic components of the learning experience onto video allows for these important instructional methods (e.g., class discussion) to take up more of the precious face-to-face time with students.

Video Is Best When Added to Existing Content

The results of this meta-analysis strongly support offering students video in addition to other classes: providing students with both learning opportunities led to significant benefits compared with just face-to-face classes. While some academics may fear that students will choose the online lecture in exchange for coming to class, our results suggest that offering both learning opportunities is very useful. Adding videos to existing content was excellent for student learning, with strong effect sizes robust to most moderation and sensitivity analyses. These effects were moderated by the funding provided to the project. If we assume funded projects allow for higher quality multimedia than unfunded ones, then this moderator may support the aforementioned discussion about the importance of quality multimedia design. Regardless of the funding, however, offering both videos and face-to-face classes would be consistent with recommendations from other meta-analyses that indicated online learning was most effective when blended with face-to-face classes (Means et al., 2009, 2010; Means et al., 2013).

Limitations of the Included Studies

None of the studies in our review met all the criteria for internal validity recommended by the *Cochrane Handbook*. Each of these criteria are important because failing to meet them has been shown to independently inflate effect sizes in meta-meta-analyses (Higgins et al., 2011). For example, studies that prospectively registered their methods and primary outcome are much less likely to find significant results (Kaplan & Irvin, 2015). So, while we optimized for causal inferences by only including randomized trials, these biases do undermine our ability to make causal inferences. Some of the Cochrane criteria are difficult to meet. Blinding students to hypotheses can be a difficult task, usually requiring either deception or obfuscation during the informed consent process. In other educational reviews that used the Cochrane Risk of Bias tool, reviewers have either failed to find any studies that were low risk of bias for participant blinding (Chung et al., 2017; Hu et al., 2018; Qin et al., 2016; S. Wood & Mayo-Wilson, 2012) or omitted this criteria because it was too difficult a standard (Kyaw et al., 2019). Blinding may influence subjective outcomes like student preferences, but less frequently influence objective outcomes like skills assessments (L. Wood et al., 2008). Therefore, it is possible that these criteria are not as important for research on learning in higher education. This argument is supported by our sensitivity analyses that showed no risk of bias criteria that moderated effects. Nevertheless, without addressing these criteria, it is more difficult for educational researchers to establish a causal link that cannot be explained by performance bias.

Where blinding students and teachers might be a challenging criteria, other criteria are easy to meet. Prospective registration is quick and free via the Open Science Framework. Adequate, transparent generation of a random sequence can be done using many software packages. Reporting of these details is easy to ensure (so that criteria are less frequently judged as “unclear”) by journals recommending that authors use reporting checklists (e.g., CONSORT; Schulz et al., 2010). Having said that, the effect sizes in our review were not significantly different when controlling for these criteria, so results appeared somewhat robust to these threats to internal validity.

The same is true about the possibility of publication bias. Results in this review were consistent with publication bias because significant studies were more likely to be published, controlling for their effect sizes. The field can control for these publication biases by publishing well-designed, adequately powered studies, regardless of their significance. Alternatively, it could promote the practice of registered reports, where articles are provisionally accepted by journals before the results are known (Nosek & Lakens, 2014). Nevertheless, while publication bias is present in many fields, the bias is generally modest, with significant results less than two times as likely to be published (Mathur & VanderWeele, 2019b). In our review, to nullify many results, studies would need to be 10 times more likely to be published if significant. As a result, while publication bias is a threat to the internal validity of the conclusions in this review, the results appear robust to this bias.

Limitations of This Review

To assess a homogeneous outcome variable, our review only assessed student learning as an outcome. This restriction necessarily excluded other potentially important outcomes that academics may hope students achieve (Warnick & Burbules, 2007). For example, we excluded studies that only reported student satisfaction, retention, engagement, or interest. Many of these variables, like student engagement, are important predictors of long-term success (Lawson & Lawson, 2013). Student satisfaction scores are often an important consideration of higher education institutions (Spooren et al., 2013). But given these scores are not highly correlated with student learning (Uttl et al., 2017), our review does not allow us to make inferences about whether or not students prefer to learn via video, or whether any of these other outcomes are maintained. Other reviews have found similar levels of student satisfaction from online learning more generally (Cook et al., 2008), but future reviews could better assess the effects of video in particular.

Similarly, by focusing on video as the only independent variable, we excluded studies where videos were part of a complex change to learning design. As mentioned earlier, videos are typically used to enable other types of instruction (Warnick & Burbules, 2007), such as by moving lectures to video to be watched before class, so face-to-face time can be focused on interactivity (Abeysekera & Dawson, 2015). By looking at the effect of video in isolation, our review may not reflect the influence of video once controlling for parallel changes in both student and teacher behavior (Paulus et al., 2012; Tallent-Runnels et al., 2006; Warnick & Burbules, 2007). Students often change their behavior when they start learning via a different method (Tallent-Runnels et al., 2006). On one hand, flexible learning from videos may increase engagement because it removes some barriers. On the other hand, the accountability imposed by face-to-face classes may mean that shifting content to video may lead to reduced student engagement. Given we only looked at randomized trials, this disengagement may be less obvious because (a) students who are disengaged do not enroll in the study or (b) students are more likely to engage when being more closely scrutinized for a research project. Frequent class attendance has strong associations with learning (Credé et al., 2010), and perhaps, being in a certain place at a certain time provides useful structure to a learning experience.

This is one example of where poor implementation of video could lead to laissez-faire teaching. Instead, faculty should ensure that video interventions are provided in a context of both autonomy and structure (e.g., using frequent formative assessment or monitoring of online engagement), given both are important in higher education (Aelterman et al., 2019; Jang et al., 2010). With this structure, some students may thrive in fully online courses (Tallent-Runnels et al., 2006) but many students may not, including those with poorer IT (information technology) literacy, or those from lower socioeconomic or educational backgrounds (Eynon & Helsper, 2011). For many students, online learning is likely to be better if blended with face-to-face classes (Means et al., 2010). Meta-analyses of online learning consistently suggest that asynchronous online learning is best when supported by synchronous activities where students can dynamically interact with teachers and their peers (Means et al., 2010). Supplementing student–video interactions with other types of interactivity (i.e., student–teacher, student–student; Bernard et al., 2009) is likely to improve learning (Chi & Wylie, 2014), but our review cannot quantify these effects because these designs involve more than one independent variable.

Our review is also limited by a substantial amount of unexplained heterogeneity. While we explained some of this heterogeneity through moderation analyses, much of the unexplained heterogeneity may be from variations in how the videos were designed. There are a number of well-established design principles drawn from the cognitive theory of multimedia learning that have been shown to increase learning (Mayer, 2008), like the decision to highlight important material (i.e., the signaling effect; S. Schneider, Beege, et al., 2018). Multimedia are more effective when broken into learner-paced chunks (Rey et al., 2019), when keywords are presented on screen (Adesope & Nesbit, 2012), when media are visually appealing (Brom et al., 2018), when language is conversational (Ginns et al., 2013), and when visuals and speech are close together, both in time and in space (Ginns, 2006).

It is likely that studies in our review varied in their use of these design principles and that these design principles account for some of the learning benefits. There may be a selection bias where educators who run randomized trials on the effectiveness of video may be above-average on their use of good multimedia design. As a result, those developing videos for learning would benefit from considering these principles if they want to replicate the benefits found in our review. However, it is also unlikely that the same benefits would apply to videos that ignored good design: students are unlikely to learn effectively from a complex, 2-hour soliloquy filled with jargon. So as with computer-based instruction, teaching staff may need training and technical support to deliver high-quality learning (Tallent-Runnels et al., 2006), particularly to shoot, edit, and postproduce well-designed content. As reviewers, we were unable to explore the influence of these design principles because studies rarely reported what design principles they used in their videos or the comparison condition. As a result, we cannot confirm whether these design principles are what explain heterogeneity in effects, or whether it is better explained by other moderators that we did not assess. So while this review shows that, on average, video increases learning, and increases are likely moderated by the multimedia design principles, future research could help identify which design principles are most influential.

Some authors recommend searching trial registries for studies (Shea et al., 2017), and a limitation of our review was that we instead focused on database and citation searching. We did not search trial registries because we judged that the likelihood of finding included studies that were not found by other methods was low. We limited our search to studies published in English for logistical reasons, because our team had neither the skills nor the funding to translate articles from other languages, so results may not generalize to non-English-speaking countries. Along these lines, the region in which the study was conducted was a significant moderator of effects when adding videos to classes (but not when swapping content for videos). This may reflect regional differences in either the familiarity or the novelty of video as a pedagogical tool, or differential access to technology (Valenzuela-Levi, 2020; Warschauer & Matuchniak, 2010). Alternatively, it may be a statistical artifact with a small number of studies in many regions. We did not prospectively register this moderator so it should be interpreted with caution, but future reviews that explore mediators of effects (e.g., if increases in novelty or access explain increases in learning) may help explain findings.

Conclusion

Online teaching allows for learning to be delivered affordably, at scale, and with fewer infrastructure constraints than face-to-face instruction—it does not require a university to have hundreds of people in the same room at the same time. As universities move content online, staff usually turn to teaching via asynchronous videos and synchronous videoconferences (Cook et al., 2010). Videoconferences may be more conducive to student–teacher interactivity (Al-Samarraie, 2019; Bernard et al., 2009), but most studies in our review had the same level of active learning in the video and comparison condition, meaning teachers can maintain active learning while shifting to video. When they do, videos lead to better student learning than many other teaching methods, even when compared with face-to-face teaching. We suggest that these results are because videos may provide students with control over their level of cognitive load, they allow authentic demonstrations of skills, and they enable teaching staff to edit according to multimedia learning principles. Pragmatically, videos allow students to fit learning around their other commitments, and are less reliant than videoconferences on stable, high-speed internet connections (Al-Samarraie, 2019), because they can be buffered to the user's device. Many of these findings are still contingent on students having access to online learning (Warschauer & Matuchniak, 2010), so ensuring students have social and technical support is a critical challenge for universities around the world. Provided those supports are available, universities can effectively switch to video for efficient and scalable education. Video appears to have a range of benefits in higher education settings.

Acknowledgments

We thank Leslie Silk who completed initial searches, screening, and extraction.

Funding

This review was partially funded by a Teaching Development Grant provided by our university.

ORCID iDs

Michael Noetel https://orcid.org/0000-0002-6563-8203
Oscar Delaney https://orcid.org/0000-0002-8792-0744
Philip Parker https://orcid.org/0000-0002-4604-8566
Borja del Pozo Cruz https://orcid.org/0000-0003-3944-2212
Chris Lonsdale https://orcid.org/0000-0002-2523-5565

References

Abeysekera, L., & Dawson, P. (2015). Motivation and cognitive load in the flipped classroom: Definition, rationale and a call for research. *Higher Education Research & Development*, *34*(1), 1–14. https://doi.org/10.1080/07294360.2014.934336

Adesope, O. O., & Nesbit, J. C. (2012). Verbal redundancy in multimedia learning environments: A meta-analysis. *Journal of Educational Psychology*, *104*(1), 250–263. https://doi.org/10.1037/a0026147

Aelterman, N., Vansteenkiste, M., Haerens, L., Soenens, B., Fontaine, J. R. J., & Reeve, J. (2019). Toward an integrative and fine-grained insight in motivating and demotivating teaching styles: The merits of a circumplex approach. *Journal of Educational Psychology*, *111*(3), 497–521. https://doi.org/10.1037/edu0000293

Albalawi, A. S. (2018). The effect of using flipped classroom in teaching calculus on students' achievements at University of Tabuk. *International Journal of Research in Education and Science*, *4*(1), 198–207. https://eric.ed.gov/?id=EJ1169830

Alexander, P. A. (2020). Methodological guidance paper: The art and science of quality systematic reviews. *Review of Educational Research*, *90*(1), 6–23. https://doi.org/10.3102/0034654319854352

Al-Samarraie, H. (2019). A scoping review of videoconferencing systems in higher education: Learning paradigms, opportunities, and challenges. *International Review of Research in Open and Distributed Learning*, *20*(3). https://doi.org/10.19173/irrodl.v20i4.4037

APA Publications and Communications Board Working Group on Journal Article Reporting Standards. (2008). Reporting standards for research in psychology: Why do we need them? What might they be? *The American Psychologist*, *63*(9), 839–851. https://doi.org/10.1037/0003-066X.63.9.839

Baxter, P., Akhtar-Danesh, N., Landeen, J., & Norman, G. (2012). Teaching critical management skills to senior nursing students: Videotaped or interactive hands-on instruction? *Nursing Education Perspectives*, *33*(2), 106–110. https://doi.org/10.5480/1536-5026-33.2.106

Bernard, R. M., Abrami, P. C., Borokhovski, E., Wade, C. 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. https://doi.org/10.3102/0034654309333844

Beydogan, H. Ö., & Hayran, Z. (2015). The effect of multimedia-based learning on the concept learning levels and attitudes of students. *Eurasian Journal of Educational Research*, *60*, 261–280. https://eric.ed.gov/?id=EJ1076689
Biggs, J. B., & Tang, C. (2011). *Teaching for quality learning at university: What the student does*. McGraw-Hill Education.
Brom, C., Stárková, T., & D'Mello, S. K. (2018). How effective is emotional design? A meta-analysis on facial anthropomorphisms and pleasant colors during multimedia learning. *Educational Research Review*, *25*, 100–119. https://doi.org/10.1016/j.edurev.2018.09.004
Chang, M.-M., & Lin, M.-C. (2014). The effect of reflective learning e-journals on reading comprehension and communication in language learning. *Computers & Education*, *71*, 124–132. https://doi.org/10.1016/j.compedu.2013.09.023
Cheung, M. W.-L. (2014). metaSEM: an R package for meta-analysis using structural equation modeling. *Frontiers in Psychology*, *5*, 1521. https://doi.org/10.3389/fpsyg.2014.01521
Chi, M. T., Adams, J., Bogusch, E. B., Bruchok, C., Kang, S., Lancaster, M., Levy, R., Li, N., McEldoon, K. L., Stump, G. S., Wylie, R., Xu, D., & Yaghmourian, D. L. (2018). Translating the ICAP theory of cognitive engagement into practice. *Cognitive Science*, *42*(6), 1777–1832. https://doi.org/10.1111/cogs.12626
Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823
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/Sigma Theta Tau International, Honor Society of Nursing*, *9*(2), 78–87. https://doi.org/10.1111/j.1741-6787.2012.00241.x
Chung, K.-F., Chan, M.-S., Lam, Y.-Y., Lai, C. S.-Y., & Yeung, W.-F. (2017). School-based sleep education programs for short sleep duration in adolescents: A systematic review and meta-analysis. *Journal of School Health*, *87*(6), 401–408. https://doi.org/10.1111/josh.12509
Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research*, *53*(4), 445–459. https://doi.org/10.3102/00346543053004445
Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development: ETR & D*, *42*(2), 21–29. http://doi.org/10.1007/BF02299088
Conigrave, J. (2019). msemtools (Version 0.9.8) [Computer software]. https://github.com/JConigrave/msemtools
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*(8), 765–774. https://doi.org/10.1111/j.1365-2923.2010.03723.x
Cook, D. A., Levinson, A. J., Garside, 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*(10), 1181–1196. https://doi.org/10.1001/jama.300.10.1181
Credé, M., Roch, S. G., & Kieszczynka, U. M. (2010). Class attendance in college: A meta-analytic review of the relationship of class attendance with grades and student characteristics. *Review of Educational Research*, *80*(2), 272–295. https://doi.org/10.3102/0034654310362998

Deeks, J. J., Higgins, J. P. T., Altman, D. G., on behalf of the Cochrane Statistical Methods Group and the Cochrane Bias Methods Group. (2011). Analysing data and undertaking meta-analyses. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions* (Vol. 5.1.1). The Cochrane Collaboration.

Deming, D. (2020, April 9). Online learning should return to a supporting role. *The New York Times*. https://www.nytimes.com/2020/04/09/business/online-learning-virus.html

Durlak, J. A. (2009). How to select, calculate, and interpret effect sizes. *Journal of Pediatric Psychology*, *34*(9), 917–928. https://doi.org/10.1093/jpepsy/jsp004

Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press.

Eriksen, M. B., & Frandsen, T. F. (2018). The impact of patient, intervention, comparison, outcome (PICO) as a search strategy tool on literature search quality: A systematic review. *Journal of the Medical Library Association: JMLA*, *106*(4), 420–431. https://doi.org/10.5195/jmla.2018.345

Eynon, R., & Helsper, E. (2011). Adults learning online: Digital choice and/or digital exclusion? *New Media & Society*, *13*(4), 534–551. https://doi.org/10.1177/1461444810374789

Farooq, I., & Al-Jandan, B. A. (2015). Effect of video triggering during conventional lectures on final grades of dental students in an oral biology course: A two-year retrospective study. *Journal of Dental Education*, *79*(12), 1467–1470. https://www.ncbi.nlm.nih.gov/pubmed/26632302

Fernandez, A., Tan, K.-A., Knaak, S., Chew, B. H., & Ghazali, S. S. (2016). Effects of brief psychoeducational program on stigma in Malaysian pre-clinical medical students: A randomized controlled trial. *Academic Psychiatry: The Journal of the American Association of Directors of Psychiatric Residency Training and the Association for Academic Psychiatry*, *40*(6), 905–911. https://doi.org/10.1007/s40596-016-0592-1

García-Holgado, A., Marcos-Pablos, S., & García-Peñalvo, F. (2020). Guidelines for performing systematic research projects reviews. *International Journal of Interactive Multimedia and Artificial Intelligence*, *6*(2), 9. https://doi.org/10.9781/ijimai.2020.05.005

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction*, *15*(4), 313–331. https://doi.org/10.1016/j.learninstruc.2005.07.001

Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction*, *16*(6), 511–525. https://doi.org/10.1016/j.learninstruc.2006.10.001

Ginns, P., Martin, A. J., & Marsh, H. W. (2013). Designing instructional text in a conversational style: A meta-analysis. *Educational Psychology Review*, *25*(4), 445–472. https://doi.org/10.1007/s10648-013-9228-0

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational and Behavioral Statistics*, *6*(2), 107–128. https://doi.org/10.3102/10769986006002107

Hedges, L. V., & Vevea, J. L. (1996). Estimating effect size under publication bias: Small sample properties and robustness of a random effects selection model. *Journal of Educational and Behavioral Statistics*, *21*(4), 299–332. https://doi.org/10.3102/10769986021004299

Higgins, J. P. T., Altman, D. G., Sterne, J. A. C., & on behalf of the Cochrane Statistical Methods Group and the Cochrane Bias Methods Group. (2011). Assessing risk of bias in included studies. In J. P. T. Higgins & S. Green (Eds.), *Cochrane* *handbook for systematic reviews of interventions* (Vol. 5.1.1). The Cochrane Collaboration. https://handbook-5-1.cochrane.org/chapter_8/8_assessing_risk_of_bias_in_included_studies.htm

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2019). *Cochrane handbook for systematic reviews of interventions*. John Wiley. https://training.cochrane.org/cochrane-handbook-systematic-reviews-interventions

Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction*, *17*(6), 722–738. https://doi.org/10.1016/j.learninstruc.2007.09.013

Hu, R., Gao, H., Ye, Y., Ni, Z., Jiang, N., & Jiang, X. (2018). Effectiveness of flipped classrooms in Chinese baccalaureate nursing education: A meta-analysis of randomized controlled trials. *International Journal of Nursing Studies*, *79*, 94–103. https://doi.org/10.1016/j.ijnurstu.2017.11.012

Ilic, D., Nordin, R. B., Glasziou, P., Tilson, J. K., & Villanueva, E. (2015). A randomised controlled trial of a blended learning education intervention for teaching evidence-based medicine. *BMC Medical Education*, *15*(1), Article 39. https://doi.org/10.1186/s12909-015-0321-6

Jang, H., Reeve, J., & Deci, E. L. (2010). Engaging students in learning activities: It is not autonomy support or structure but autonomy support and structure. *Journal of Educational Psychology*, *102*(3), 588–600. https://doi.org/10.1037/a0019682

Johnson, H., & Mejia, M. C. (2014). *Online learning and student outcomes in California's community colleges*. Public Policy Institute of California. https://www.ppic.org/content/pubs/report/R_514HJR.pdf

Kaplan, R. M., & Irvin, V. L. (2015). Likelihood of null effects of large NHLBI clinical trials has increased over time. *PLoS One*, *10*(8), e0132382. https://doi.org/10.1371/journal.pone.0132382

Kay, R. H. (2012). Exploring the use of video podcasts in education: A comprehensive review of the literature. *Computers in Human Behavior*, *28*(3), 820–831. https://doi.org/10.1016/j.chb.2012.01.011

Kozma, R. B. (1994). Will media influence learning? Reframing the debate. *Educational Technology Research and Development: ETR & D*, *42*(2), 7–19. https://doi.org/10.1007/BF02299087

Kyaw, B. M., Saxena, N., Posadzki, P., Vseteckova, J., Nikolaou, C. K., George, P. P., Divakar, U., Masiello, I., Kononowicz, A. A., Zary, N., & Tudor Car, L. (2019). Virtual reality for health professions education: Systematic review and meta-analysis by the digital health education collaboration. *Journal of Medical Internet Research*, *21*(1), e12959. https://doi.org/10.2196/12959

Lawson, M. A., & Lawson, H. A. (2013). New conceptual frameworks for student engagement research, policy, and practice. *Review of Educational Research*, *83*(3), 432–479. https://doi.org/10.3102/0034654313480891

Lazowski, R. A., & Hulleman, C. S. (2016). Motivation interventions in education: A meta-analytic review. *Review of Educational Research*, *86*(2), 602–640. https://doi.org/10.3102/0034654315617832

Lee, K. (2020, March 9). Coronavirus: Universities are shifting classes online—but it's not as easy as it sounds. *The Conversation*. http://theconversation.com/coronavirus-universities-are-shifting-classes-online-but-its-not-as-easy-as-it-sounds-133030

Mathur, M. B., & VanderWeele, T. J. (2019a). New metrics for meta-analyses of heterogeneous effects. *Statistics in Medicine*, *38*(8), 1336–1342. https://doi.org/10.1002/sim.8057

Mathur, M. B., & VanderWeele, T. (2019b). Sensitivity analysis for publication bias in meta-analyses. In *Open Science Framework*. https://doi.org/10.31219/osf.io/s9dp6

Mayer, R. E. (2008). Applying the science of learning: Evidence-based principles for the design of multimedia instruction. *The American Psychologist*, *63*(8), 760–769. https://doi.org/10.1037/0003-066X.63.8.760

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press.

Mayer, R. E., Fiorella, L., & Stull, A. (2020). Five ways to increase the effectiveness of instructional video. *Educational Technology Research and Development: ETR & D*. https://doi.org/10.1007/s11423-020-09749-6

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*(1), 43–52. https://doi.org/10.1207/S15326985EP3801_6

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. https://www.sri.com/sites/default/files/publications/effectiveness_of_online_and_blended_learning.pdf

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

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2010). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies (revised)*. U.S. Department of Education.

Moeyaert, M., Ugille, M., Natasha Beretvas, S., Ferron, J., Bunuan, R., & Van den Noortgate, W. (2017). Methods for dealing with multiple outcomes in meta-analysis: a comparison between averaging effect sizes, robust variance estimation and multilevel meta-analysis. *International Journal of Social Research Methodology*, *20*(6), 559–572. https://doi.org/10.1080/13645579.2016.1252189

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery*, *8*(5), 336–341. https://doi.org/10.1016/j.ijsu.2010.02.007

Moher, D., Tetzlaff, J., Tricco, A. C., Sampson, M., & Altman, D. G. (2007). Epidemiology and reporting characteristics of systematic reviews. In *PLOS Medicine*, *4*(3), e78. https://doi.org/10.1371/journal.pmed.0040078

Muller, D. A., Bewes, J., Sharma, M. D., & Reimann, P. (2007). Saying the wrong thing: Improving learning with multimedia by including misconceptions. *Journal of Computer Assisted Learning*, *24*(2), 144–155. https://doi.org/10.1111/j.1365-2729.2007.00248.x

Muller, D. A., Sharma, M. D., & Reimann, P. (2008). Raising cognitive load with linear multimedia to promote conceptual change. *Science Education*, *92*(2), 278–296. https://doi.org/10.1002/sce.20244

Nosek, B. A., & Lakens, D. (2014). Registered reports. *Social Psychology*, *45*(3), 137–141. https://doi.org/10.1027/1864-9335/a000192

O'Donovan, J., Ahn, R., Nelson, B. D., Kagan, C., & Burke, T. F. (2016). Using low-cost Android tablets and instructional videos to teach clinical skills to medical students in Kenya: A prospective study. *JRSM Open*, *7*(8), 2054270416645044. https://doi.org/10.1177/2054270416645044

Paulus, T. M., Phipps, G., Harrison, J., & Varga, M. A. (2012). Re-envisioning instructional technology research in higher education environments: A content analysis of a grant program. *Journal of Computing in Higher Education*, *24*(3), 164–181. https://doi.org/10.1007/s12528-012-9062-2

Pigott, T. D., & Polanin, J. R. (2020). Methodological guidance paper: High-quality meta-analysis in a systematic review. *Review of Educational Research*, *90*(1), 24–46. https://doi.org/10.3102/0034654319877153

Pustejovsky, J. E., & Rodgers, M. A. (2019). Testing for funnel plot asymmetry of standardized mean differences. *Research Synthesis Methods*, *10*(1), 57–71. https://doi.org/10.1002/jrsm.1332

Qin, Y., Wang, Y., & Floden, R. E. (2016). The effect of problem-based learning on improvement of the medical educational environment: A systematic review and meta-analysis. *Medical Principles and Practice: International Journal of the Kuwait University, Health Science Centre*, *25*(6), 525–532. https://doi.org/10.1159/000449036

R Core Team. (2020). R: A language and environment for statistical computing (Version 3.6.3) [Computer software]. *R Foundation for Statistical Computing*. https://www.R-project.org/

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, *147*, 103778. https://doi.org/10.1016/j.compedu.2019.103778

Rath, T. E., & Holt, D. W. (2010). Vicarious audiovisual learning in perfusion education. *Journal of Extra-Corporeal Technology*, *42*(4), 305–312. https://www.ncbi.nlm.nih.gov/pubmed/21313929

Rey, G. D. (2012). A review of research and a meta-analysis of the seductive detail effect. *Educational Research Review*, *7*(3), 216–237. https://doi.org/10.1016/j.edurev.2012.05.003

Rey, G. D., Beege, M., Nebel, S., Wirzberger, M., Schmitt, T. H., & Schneider, S. (2019). A meta-analysis of the segmenting effect. *Educational Psychology Review*, *31*(2), 389–419. https://doi.org/10.1007/s10648-018-9456-4

Rolfe, V. E., & Gray, D. T. (2011). Are multimedia resources effective in life science education? A meta-analysis. *Bioscience Education*, *18*(1), 1–14. https://doi.org/10.3108/beej.18.5

Sayed, I. E., & Abdelmonem, S. (2018). 14 Evidence based medicine in low-middle income countries; online or face-to-face teaching method? *Oral Sessions*, *23*, A7.2–A7. https://doi.org/10.1136/bmjebm-2018-111024.14

Schneider, S., Beege, M., Nebel, S., & Rey, G. D. (2018). A meta-analysis of how signaling affects learning with media. *Educational Research Review*, *23*, 1–24. https://doi.org/10.1016/j.edurev.2017.11.001

Schneider, S., Nebel, S., Beege, M., & Rey, G. D. (2018). The autonomy-enhancing effects of choice on cognitive load, motivation and learning with digital media. *Learning and Instruction*, *58*, 161–172. https://doi.org/10.1016/j.learninstruc.2018.06.006

Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin*, *143*(6), 565–600. https://doi.org/10.1037/bul0000098

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

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. *British Medical Journal*, *358*, j4008. https://doi.org/10.1136/bmj.j4008

Shrivastav, H., & Hiltz, S. R. (2013). Information overload in technology-based education: A meta-analysis. *AMCIS 2013 Proceedings*. https://aisel.aisnet.org/amcis2013/ISEducation/GeneralPresentations/6/

Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research*, *83*(4), 598–642. https://doi.org/10.3102/0034654313496870

Steedman, M., Abouammoh, M., & Sharma, S. (2012). Multimedia learning tools for teaching undergraduate ophthalmology: Results of a randomized clinical study. *Canadian Journal of Ophthalmology/Journal Canadien D'ophtalmologie*, *47*(1), 66–71. https://doi.org/10.1016/j.jcjo.2011.12.006

Tallent-Runnels, M. K., Thomas, J. A., Lan, W. Y., Cooper, S., Ahern, T. C., Shaw, S. M., & Liu, X. (2006). Teaching courses online: A review of the research. *Review of Educational Research*, *76*(1), 93–135. https://doi.org/10.3102/00346543076001093

The EndNote Team (2013). EndNote (Version X9). Clarivate. https://endnote.com/

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*(2), 70–75. https://doi.org/10.1071/NB12076

Tularam, G. A., & Machisella, P. (2018). Traditional vs non-traditional teaching and learning strategies: The case of e-learning! *International Journal for Mathematics Teaching and Learning*, *19*(1), 129–158. http://www.cimt.org.uk/ijmtl/index.php/IJMTL/article/view/21

Uttl, B., Cnudde, K., & White, C. A. (2019). Conflict of interest explains the size of student evaluation of teaching and learning correlations in multisection studies: A meta-analysis. *PeerJ*, *7*, e7225. https://doi.org/10.7717/peerj.7225

Uttl, B., White, C. A., & Gonzalez, D. W. (2017). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation*, *54*, 22–42. https://doi.org/10.1016/j.stueduc.2016.08.007

Valenzuela-Levi, N. (2020). The written and unwritten rules of internet exclusion: Inequality, institutions and network disadvantage in cities of the Global South. *Information, Communication and Society*, 1–18. https://doi.org/10.1080/1369118X.2020.1716037

Van Gog, T., Ericsson, K. A., Rikers, R. M., & Paas, F. (2005). Instructional design for advanced learners: Establishing connections between the theoretical frameworks of cognitive load and deliberate practice. *Educational Technology Research and Development: ETR & D*, *53*(3), 73–81. https://doi.org/10.1007/BF02504799

van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive Load Theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, *17*(2), 147–177. https://doi.org/10.1007/s10648-005-3951-0

Veletsianos, G., & Houlden, S. (2020, March 12). Coronavirus pushes universities to switch to online classes: But are they ready? *The Conversation*. http://theconversation.com/coronavirus-pushes-universities-to-switch-to-online-classes-but-are-they-ready-132728

Viechtbauer, W. (2017). Package "metafor" (Version 2.0) [Computer software]. https://github.com/wviechtb/metafor

Warnick, B. R., & Burbules, N. C. (2007). Media comparison studies: Problems and possibilities. *Teachers College Record*, *109*(11), 2483–2510.

Warschauer, M., & Matuchniak, T. (2010). New technology and digital worlds: Analyzing evidence of equity in access, use, and outcomes. *Review of Research in Education*, *34*(1), 179–225. https://doi.org/10.3102/0091732X09349791

Wood, L., Egger, M., Gluud, L. L., Schulz, K. F., Jüni, P., Altman, D. G., Gluud, C., Martin, R. M., Wood, A. J. G., & Sterne, J. A. C. (2008). Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: Meta-epidemiological study. *BMJ*, *336*(7644), 601–605. https://doi.org/10.1136/bmj.39465.451748.AD

Wood, S., & Mayo-Wilson, E. (2012). School-based mentoring for adolescents: A systematic review and meta-analysis. *Research on Social Work Practice*, *22*(3), 257–269. https://doi.org/10.1177/1049731511430836

Authors

MICHAEL NOETEL is a senior lecturer in the School of Health and Behavioural Sciences at the Australian Catholic University, 1100 Nudgee Road, Banyo, Queensland 4014, Australia; email: *michael.noetel@acu.edu.au*. He received his doctorate in psychology from the Australian Catholic University. His major research interests include online learning, motivation, and research synthesis.

SHANTELL GRIFFITH is a postgraduate student in social work at the Australian Catholic University, 1100 Nudgee Road, Banyo, Queensland 4014, Australia; email: *shantell.griffith@myacu.edu.au*. She received her bachelor of psychological science from the Australian Catholic University.

OSCAR DELANEY is an undergraduate biology student at the University of Queensland (1/231 Fairfield Road, Fairfield, Queensland 4103, Australia; email: *o.delaney@uq.net.au*) with research interests in climate adaptation and plant genetics.

TAREN SANDERS is the Deputy Program Leader for the Motivation and Behaviour Research Program in the Institute for Positive Psychology and Education at the Australian Catholic University, 1100 Nudgee Road, Banyo, Queensland 4014, Australia; email: *taren.sanders@acu.edu.au*. He received his doctorate from Western Sydney University. His major research interest includes screen use and physical activity of children and young people.

PHILIP PARKER is the Deputy Director of the Institute for Positive Psychology and Education at the Australian Catholic University, 1100 Nudgee Road, Banyo, Queensland 4014, Australia; email: *philip.parker@acu.edu.au*. He received his doctorate in educational psychology from the University of Sydney. His major research interest includes educational inequality, developmental transitions, and educational attainment.

BORJA DEL POZO CRUZ is Senior Research Fellow of the Institute for Positive Psychology and Education at the Australian Catholic University, 1100 Nudgee Road, Banyo, Queensland 4014, Australia; email: *Borja.delPozoCruz@acu.edu.au*. He received his doctorate from the University of Pablo de Olavide. His major research interest includes health promotion, physical activity, and analytical methods.

CHRIS LONSDALE is the Program Leader for the Motivation and Behaviour Research Program in the Institute for Positive Psychology and Education at the Australian Catholic University, 1100 Nudgee Road, Banyo, Queensland 4014, Australia; email: *chris.lonsdale@acu.edu.au*. He received his doctorate from the University of Otago. His major research interest includes the motivation underpinnings of behavior in education and health contexts.