



# Instructors' presence in instructional videos: A systematic review

Hamza Polat<sup>1</sup>

Received: 2 October 2022 / Accepted: 13 December 2022 / Published online: 23 December 2022  
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

## Abstract

The discussion about how to present instructors in instructional videos has become a hot topic in recent years. This systematic review explores how the instructors' presence affects affective, cognitive, and social aspects of learning in different conditions and with different video types. The review analyses 41 empirical studies indexed in Web of Science, ERIC, Scopus, and Education Source research databases from 2014 to 2022. The results indicated that (i) many instructor-present videos were in picture-in-picture format and included direct gaze as a social cue, (ii) learners had positive feelings for instructor-present videos, (iii) the on-screen instructor could not be beneficial for gathering positive learning outcomes, but social and attentional cues provided by the on-screen instructor could foster learning, and (iv) findings regarding the social aspect of learning were inconclusive. This study also emphasizes the need for further studies to clearly explore the role of the instructor in different learning conditions.

**Keywords** Instructor's presence · Instructional video · Learning outputs · Video design · Systemic review

## 1 Introduction

Videos have become more prevalent in delivering learning content in online instruction since the emergence of the COVID-19 pandemic. With the widespread use of videos in education, numerous design issues have been applied to make videos more appealing, engaging (Wilson et al., 2018) and effective for instruction (van Wermeskerken & van Gog, 2017). However, the intrinsic transactional distance generated by online environments is still a challenging problem for instruction (Wang

---

✉ Hamza Polat  
hamzapolat@atauni.edu.tr

<sup>1</sup> Department of Information Systems and Technologies, Ataturk University, 25240 Erzurum, Turkey

et al., 2020b). A commonly emphasized design issue to overcome this problem and improve students' engagement and learning performance is the on-screen presence of instructors in videos (van Wermeskerken & van Gog, 2017).

The literature suggests two general classifications to categorize instructor-presence videos (IPV): lecture and demonstration videos (Stull et al., 2018; van Wermeskerken et al., 2018). In lecture-style videos, the instructor explains a material by standing next to a board on which learning content (i.e., figures, drawings, etc.) is presented or a screen on which the slides are projected to audiences; in demonstration-style videos, on the other hand, the instructor demonstrates a learning task (van Wermeskerken & van Gog, 2017). The instructor might be either in full-body or talking-head forms (Stull et al., 2018) and is regarded as a social cue (Yi et al., 2019), who use facial or body expressions, gestures, and hand moves, or provides particular attentional cues.

Although various studies have examined the effect of the instructor's presence on learning outcomes, the results are inconclusive and mixed (Wang et al., 2020a). While some studies have shown that learning from IPVs is more effective than learning from those without the instructor (Pi & Hong, 2016), others have found that learning from IPVs is ineffective (Kizilcec et al., 2014) or distracting for learners (van Wermeskerken et al., 2018). In addition to the cognitive learning outputs, there have been studies focusing various social and affective learning outcomes on IPVs, including emotion (Beege et al., 2020), satisfaction (Zhang et al., 2021), learning enjoyment (Wilson et al., 2018), self-efficacy (Hoogerheide et al., 2018), social presence (Ng & Przybylek, 2021), and parasocial interaction (Pi et al., 2021). However, it is unclear whether or not the instructor's presence is an effective strategy for these outputs since the research has not been thoroughly synthesized (Henderson & Schroeder, 2021). Given the crucial role of both social and affective processes in digital learning environments (Schneider et al., 2022), it might be beneficial to review IPV literature from different learning aspects to better understand how these processes differ under various conditions.

## 2 Prior reviews, meta-analyses, and the present study

The instructor's presence has been discussed in the literature; however, few reviews have guided the new research. The review conducted by Henderson and Schroeder (2021) examined the effect of on-screen instructors on learning, cognitive load, and social presence. They synthesized 12 peer-reviewed articles published from 2011 to 2021 and revealed descriptive trends of these videos (e.g., topic, year, and location). The review also reported neither positive findings in favor of the instructor's presence nor convincing evidence for leaving instructors out of the videos for learning outcomes, cognitive load, and social presence. Similarly, Alemdag (2022) conducted a meta-analysis with 20 experimental studies to clarify the overall impact of IPVs on learning, cognitive load, motivation, and social presence. The study showed that the instructor's presence did not influence learning and social presence but increased learners' cognitive load and motivation. Additionally, the study focused on several moderators, including video length, learning domain, human embodiment, study

setting, and study location. Findings indicated that the instructor's hand in the videos and the setting where learners watch the videos increase the impact of IPVs.

Although the abovementioned studies have provided different perspectives, the literature lacks studies describing the link between on-screen instructors and different aspects of learning. In other words, the previous works (Alemdag, 2022; Henderson & Schroeder, 2021) focused on the variables such as learning, cognitive load, social presence, and motivation and revealed some descriptive trends and moderators for IPVs. However, the current study extends earlier reviews by including new affective, cognitive, and social learning outputs, and it seeks the design-related trends in instructor-present videos. First, it divides the learning outputs into three groups, which are retention, transfer, and overall learning performance. This is important because multimedia learning literature suggests focusing on learners' retention and transfer rather than learning performance alone (Mayer, 2005). Based on empirical results in the literature, the current study may contribute to the existing knowledge of how instructors' presence affects learners' retention, transfer, and learning performance separately.

Second, the current study examines learners' attention in the IPV literature, which is also missing in previous reviews. Attentional processes can play an important role in designing multimedia learning materials (Mayer, 2014a). It is known that the learner's information processing behavior is greatly affected by the visual information on the screen (Moon & Ryu, 2020). Several eye-tracking studies, for example, have shown that the instructor's presence considerably impacts the learner's visual attention (Kizilcec et al., 2014; van Wermeskerken et al., 2018), which may distract learner from the learning content since they use the working memory resources to process the visual features of the instructor (Fiorella et al., 2019). In this regard, revealing how learners draw their attention to instructor-present videos from the literature may give insights into the design of instructional videos.

Third, the extensions of cognitive learning theories emphasize that learning performance depends not only on the cognitive processing of information but also on affective and motivational factors (Schneider et al., 2022). The current study, therefore, reviews the learners' affective states in IPV literature, which were not given much attention in the previous reviews. As an affective variable, we focused on the learners' emotions, satisfaction, and enjoyment. These variables were selected in the current review to seek learners' subjective reactions to the IPVs in the literature. It is considered that the result of this review may have a valuable contribution to the field by clarifying the role of IPVs on the affective aspect of learning.

Forth, the social cues provided by the instructor in videos lead to a sense of social presence in the learner, resulting in increased cognitive processing and more profound learning performance (Mayer, 2014b). Although previous reviews touched upon the learners' social presence, little attention is paid to the parasocial interaction, a critical variable that researcher frequently discussed in the multimedia learning environments recently (e.g., Pi et al., 2021). Parasocial interaction is described as the affective, cognitive, and behavioral influence of the media figures (e.g., instructor in the video) on the media recipient (e.g., learner) and is worth examining since it can positively impact cognitive processing (Schneider et al., 2022). Understanding both learners' social presence and parasocial processes in IPV settings may give insights into the design of instructional videos from a social

learning dimension. In this regard, this study extends the findings of previous reviews regarding learners' social presence and allows examining how the instructor's presence affects parasocial interaction depending on the literature.

Last, the design-related issues of videos, other than the visual of the instructor, may also help to interpret the findings in IPV literature. Alemdag (2022) proposed five potential moderators for knowledge acquisition from IPVs: video length, domain, human embodiment, study setting, and study location. Apart from these variables, however, there are also additional factors that are manipulated in IPVs, such as knowledge type presented in the video (e.g., Hong et al., 2018), video style (e.g., lecture or demonstration; Hew & Lo, 2020), and social cues used by the instructor (e.g., Pi et al., 2021). Classifying these factors with their varieties may reveal the trends in the design of IPVs and guide further research. Therefore, the current study focuses on categorizing design-related issues of IPVs, including video length, knowledge type, video style, and social cues.

Consequently, to get a holistic picture of diverse learning outcomes from IPVs, it remains an open question concerning under which conditions and with which video type students could attain more positive learning outcomes from instructor-present videos (Wang et al., 2020a). The current study aims to systematically review the recent literature that explores how instructors' presence affects affective, cognitive, and social aspects of learning in different conditions and with different video types. Furthermore, it also extends the results of previous works (Alemdag, 2022; HENDERSON & SCHROEDER, 2021) and addresses the following questions.

1. What is the influence of instructor-present videos on affective, cognitive, and social learning outcomes?
2. What are the design-related issues of instructor-present videos?

### 3 Method

A systematic literature review was conducted to review the recent research on instructor-present videos. This method allows rigorous and reliable knowledge organization (Denyer & Tranfield, 2009) and projects particular insights into the field through theoretical synthesis (Tranfield et al., 2003). It may be *author-centric*, which uses a chronological review to uncover the causes of a problem, or *theme-centric*, which shows readers how prior research has advanced understanding of themes and phenomena of interest (Linnenluecke et al., 2019). This study used the theme-centric approach based on the PRISMA principles (Moher et al., 2009).

#### 3.1 Search strategy

The study included three steps to find the relevant studies respectively: (i) search in the Web of Science, ERIC, Scopus, and Education Source digital research databases in January 2021, (ii) an additional search in the same databases in July 2022, and (iii) final search in Google Scholar to access missing articles. The following search string was used in all fields of each database: (“instructor\* face” OR “instructor\* gaze” OR “instructor\* head”

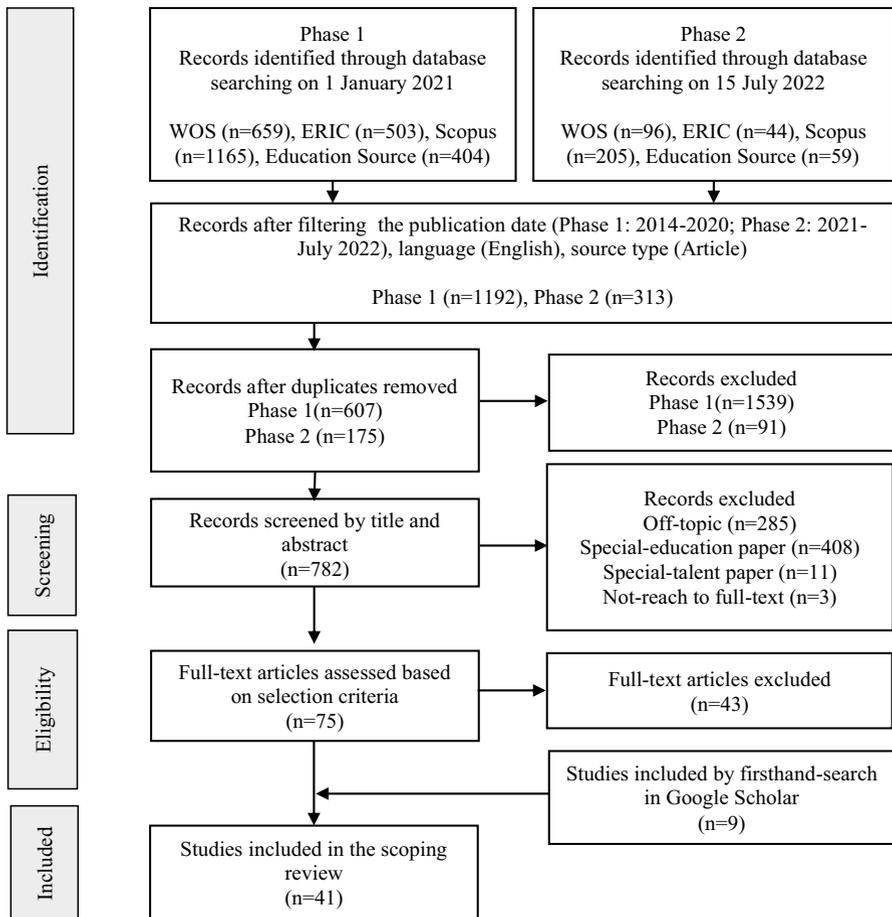
**Table 1** Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
Peer-reviewed empirical research in full text	Proceedings, dissertations, or book chapters
At least one video should include the instructor’s presence (either face or body)	Video modeling examples for individuals with special needs
Focused on educational outcomes	Not written in English
Written in English	

OR “instructor\* presence” OR “video modeling” OR “face in video” OR “social cue”) AND (“video”). The inclusion and exclusion criteria are presented in Table 1.

### 3.2 Study selection

As seen in Fig. 1, the initial search yielded 2731 articles at the beginning of 2021. The records were filtered based on publication date, language, and source type, and



**Fig. 1** PRISMA flow diagram for the study

**Table 2** Categories for synthesis

Categories	Research question
Research themes	RQ1
Affective (emotions, satisfaction, and enjoyment)	
Cognitive (attention, learning, retention, transfer, cognitive load)	
Social (social presence and parasocial interaction)	
Design-related issues	RQ2
Video length	
Knowledge type	
Video style	
Cues	

we reached a total of 1192 articles, 1539 of which were excluded because of duplicates. The remaining 607 articles' titles and abstracts were examined, and irrelevant articles were excluded from the list, resulting in the selection of 22 relevant articles. The same procedure was applied for the second search, and 10 studies were reached. We also obtained nine articles from an additional search in Google Scholar. Consequently, 41 articles were taken into consideration.

### 3.3 Identifications of categories and synthesis

The articles were examined depending on the categories presented in Table 2. The first category (research themes) was synthesized by mostly studying the articles' results section. The last category (design-related issues) was captured by examining the articles' design section. A descriptive analysis was undertaken to reflect the rates of categories, mainly frequencies and cross tables. The list of the reviewed articles and the categories for the synthesis were presented in Appendix 1.

## 4 Results

### 4.1 Research themes

#### 4.1.1 Affective learning outputs

**Emotion** Students had positive feelings towards instructor-present videos (Kizilcec et al., 2015; Wang et al., 2020a; Yuan et al., 2021), but not as much as in face-to-face instruction (Rosenthal & Walker, 2020), and showed stronger interpersonal attraction to human instructors than robot instructors on the screen (Li et al., 2016). Additionally, video style (Chen & Wu, 2015) and gestures (i.e., beat and deictic gestures; Beege et al., 2020) were not effective on learners' emotions.

**Satisfaction** Without considering the content difficulty (Wang et al., 2020a), learners' satisfaction was higher while watching instructor-present videos (Wang & Antonenko, 2017), particularly when the instructor was presented on the right side of the screen (Zhang et al., 2021). Compared to the videos with the continuous instructor presentation, strategic presentation led to more learner satisfaction (Yi et al., 2019), but video style did not (Alasfor, 2021; Yuan et al., 2021).

**Learning enjoyment** Instructor-present videos were more enjoyable than voiceover videos (Wilson et al., 2018), particularly learning from a male instructor was more enjoyable for male learners than females (Hoogerheide, Loyens, et al., 2016a). However, the instructor's age, experience (Hoogerheide et al., 2016b), and gender (Hoogerheide et al., 2018) did not influence learning enjoyment.

#### 4.1.2 Cognitive learning outputs

**Attention** Except for the articles using questionnaires (Chen & Wu, 2015; Korving et al., 2016; Rosenthal & Walker, 2020) and EEG headsets (Lackmann et al., 2021), all the articles employed eye measurements to infer the learners' attention. Additionally, learners' visual attention distribution and fixation behaviors were measured with different eye tracking metrics, including fixation duration and count, saccades, first fixation time, and dwell time.

Some articles examined how learners' attention changed in different video types, and mixed results were reported. The learners' attention in the live composite video style was greater than in picture-in-picture and voiceover videos (Rosenthal & Walker, 2020). Sustained attention in the voiceover was higher than in picture-in-picture (Chen & Wu, 2015). In another study (Korving et al., 2016), however, picture-in-picture and voiceover conditions did not significantly differ in learners' attention in the first watching of videos, but there was a difference in the second watch. The info-graphic-supported voiceover videos received more attention than lecture videos (Lackmann et al., 2021).

The instructor's face (Ouweland et al., 2015; Wang et al., 2020a; Wang & Antonenko, 2017) and the instructor with teaching enthusiasm (Qian et al., 2022) drew learners' attention more to instructor, particularly in easy-topic videos (Wang & Antonenko, 2017), which decreased over time when they rewatched the videos (van Wermeskerken et al., 2018). The most attractive stimuli that caused learners to focus on the instructor's face were their gaze behaviors, such as direct gaze (Pi et al., 2020) and guided gaze in demonstration videos (van Wermeskerken & van Gog, 2017). The other component was the whiteboard type used in video lectures. For instance, learners paid more attention to the instructor's face when watching a video lecture with a transparent whiteboard than a conventional whiteboard (Stull et al., 2021). However, when the instructor used direct gaze, the whiteboard type did not affect the learner's attention (Stull et al., 2018).

The guided gaze (Pi et al., 2020; Stull et al., 2021; Wang et al., 2018), pointing gesture (Pi et al., 2019), and gesture cues (Ouweland et al., 2015) were superior at

drawing attention to the learning content. Regardless of the instructor's gaze behavior, learners focused more on the learning material in the traditional whiteboard condition than transparent whiteboard condition (Stull et al., 2018). In demonstration videos, learners paid more attention to the learning tasks in cases where the instructor used guided gaze (van Wermeskerken & van Gog, 2017), or the instructor was not presented (van Wermeskerken et al., 2018).

The instructor's presence did not affect the fixation on learning content in picture-in-picture (Colliot & Jamet, 2018) and demonstration videos (Van Gog et al., 2014). However, when the instructor was on the right side of the screen (Zhang et al., 2021) and showed social cues, learners fixed more on the learning content. Additionally, the initial fixation on the learning content was longer, and the dwell time was shorter when the instructor used guided gaze with a surprised face (Pi et al., 2021). Furthermore, when the instructor showed direct gaze, learners focused more on the instructor's face and less on the learning content (Pi et al., 2022a).

**Learning performance** The results for learning performance were inconclusive. Picture-in-picture and video lectures was more effective than voiceover videos (Chen & Wu, 2015). This result was extended by Kokoç et al. (2020), indicating that learners' sustained attention could influence learning performance when studying videos in different styles. Additionally, the right of the screen was the most effective position to present the instructor on the screen (Zhang et al., 2021). However, some studies did not find meaningful difference between different video types (Alasfor, 2021; Wilson et al., 2018, Experiment 2,3 & 4). On the other hand, Lackmann et al. (2021) found the voiceover videos including infographics more effective than lecture videos. Additionally, the instructor's presentation (i.e., either static or strategic) style (Kizilcec et al., 2015), and whiteboard type in lecture video (Stull et al., 2018) did not affect the learning performance.

Regarding how to present the learning content, the dynamic (drawing) presentation provided more promising results than the static one, particularly in transparent whiteboard conditions (Fiorella et al., 2019). For declarative video contents, both the instructor's presence (Hong et al., 2018) and gaze behavior (Wang et al., 2018) had an impact on learning performance, while only the gaze behavior was superior for procedural knowledge (Wang et al., 2018).

Neither the gaze nor the gestures facilitated learners' performance in a problem-solving task video (Ouweland et al., 2015). However, the instructor's teaching enthusiasm (Qian et al., 2022), the instructor's pointing gesture in videos with complex contents (Pi et al., 2022b), gaze and facial expressions, particularly the interaction of direct gaze and happy face, influenced the learning performance (Pi et al., 2022a). The learning performance became weaker when the instructor performed guided gaze with a surprised face (Pi et al., 2021). Contrary to the pointing gesture condition, learners performed higher in the beat and depictive gesture condition (Pi et al., 2022b).

In the demonstration videos, the instructor's gender (Hoogerheide et al., 2018; Hoogerheide et al., 2016a) and experience (Hoogerheide et al., 2016b) did not affect the learning performance, but the age did (Hoogerheide et al., 2016b). The instructor's presence would be effective in learning performance in the second watch (Van Gog et al., 2014) as opposed to the first watch (Van Gog et al., 2014; van Wermeskerken & van Gog, 2017).

**Retention** Learners' retention was examined from different perspectives, including the instructor's presence, content's complexity, video type, gaze behavior, instructors' body orientation, gestures, and handwriting. Results were inconclusive in general. While Colliot and Jamet (2018) found videos with the instructor more effective than those without the instructor, particularly when the instructor with a professional coherence (Beege et al., 2022), others (Ng & Przybylek, 2021; van Wermeskerken et al., 2018; Wang et al., 2020b) did not observe any differences, and Yuan et al. (2021) reported more positive results in favor of the instructor's absence.

The instructor's presence did not affect learners' retention in videos covering complex topics (Wang et al., 2020a; Wang & Antonenko, 2017). However, mixed results were reported for easy topics; one study (Wang & Antonenko, 2017) found the instructor superior in retention, while another study (Wang et al., 2020a) presented non-significant findings.

In videos comparing the effect of video style on retention, no main difference was discovered (Hew & Lo, 2020; Rosenthal & Walker, 2020, Study 2). However, the way how the instructor is presented (e.g., continuous or intermitted; Yi et al., 2019), the whiteboard type (Fiorella et al., 2019, Experiment 2), and the instructor's nature (e.g., human or animated pedagogical agent; Li et al., 2016) were the possible conditions that may affect the retention.

For gaze behavior, the guided gaze was more effective in retention than the direct and averted gaze (Pi et al., 2020). Furthermore, the direct gaze was more influential than the averted gaze (Pi et al., 2020), particularly when combined with the pointing gesture (Pi et al., 2019).

The results were inconclusive for the pointing gesture condition. Pi et al. (2019) reported a significant effect of the pointing gesture on retention compared to no social cue condition, but, Fiorella and Mayer (2016) found no difference (Experiment 1). Again, no significant difference was found between the videos, including pointing gestures and hand drawings (Fiorella & Mayer, 2016). However, deictic gestures were superior for retention (Beege et al., 2020).

Instructor-present videos containing drawings were reported to be more effective on retention than videos containing static information presentation (Fiorella et al., 2019, Experiment 1). However, neither the hand nor the instructor's body significantly affected retention during the drawing (Fiorella & Mayer, 2016, Experiment 4). Similar results were obtained while drawing on a transparent whiteboard compared to a conventional whiteboard (Fiorella et al., 2019, Study 3).

Results regarding the instructor's body orientation were likewise inconsistent. Beege et al. (2017, 2019) found the frontal body orientation more effective than lateral orientation in lecture videos; however, Pi et al. (2020) did not find a difference in picture-in-picture videos. Additionally, the instructor's proximity to the camera while lecturing did not impact learners' retention scores (Beege et al., 2017).

**Transfer** The instructor's presence in demonstration videos did not affect the learners' transfer scores (Van Gog et al., 2014; van Wermeskerken et al., 2018). Conflicting results, however, were obtained for picture-in-picture and voiceover video lectures. While Wang et al. (2020b) reported positive results favoring the instructor's presence, some researchers (Colliot & Jamet, 2018; Hew & Lo, 2020, Study 2a) did not observe any differences, and Yuan et al. (2021) found voiceover videos more promising. Some studies examined how the content difficulty level affected transfer in instructor-present videos. The results were again inconclusive, indicating no effect, regardless of the content difficulty (Wang & Antonenko, 2017), but influential in videos with complex content (Wang et al., 2020a).

The instructor's intermitted (i.e., a strategic view) presence (Yi et al., 2019) and professional coherence (Beege et al., 2022) was beneficial for transfer. However, the instructor's gender (Hoogerheide et al., 2016a), proximity to the camera (Beege et al., 2017), and content presentation (i.e., static or drawing content) style and whiteboard type (Fiorella et al., 2019) had no meaningful effect on transfer scores. The results for body orientation were mixed; some (Beege et al., 2017; Van Gog et al., 2014) reported non-significant results, but Beege et al. (2019) found frontal-style more beneficial than lateral condition.

Some of the social cues shown by the instructors, such as guided gaze (Pi et al., 2020), deictic gestures (Beege et al., 2020, Experiment 2), and pointing gestures with direct gaze (Pi et al., 2019), were fruitful in getting better transfer scores. However, making direct gaze alone was not effective (Pi et al., 2019). It was also indicated that using gaze together with gestures does not necessarily increase the transfer scores (Ouweland et al., 2015). Additionally, learners' prior knowledge was noted as a crucial factor influencing the learning transfer (Fiorella & Mayer, 2016).

**Cognitive load** Mixed results were reported regarding the participants' cognitive load in the reviewed articles. No difference on cognitive load was observed between picture-in-picture and voiceover conditions (Ng & Przybylek, 2021). However, Kizilcec et al. (2015) showed that learners' cognitive load in instructor-present videos was fewer than those without the instructor (Phase 1). While this result was confirmed for intrinsic and extraneous load, it was not affirmed for the germane and overall load (Wang et al., 2020b). A mediating effect of the learning style was also noticed on the cognitive load (Chen & Wu, 2015; Kizilcec et al., 2015, Phase 2).

Some articles indicated that the knowledge type, the content's complexity, teaching enthusiasm, and instructor's professional congruence might affect the cognitive load. For example, Hong et al. (2018) discovered that the instructor's face could not affect the cognitive load in videos, including declarative knowledge (Phase 1) but could increase it in videos with procedural knowledge (Phase 2). Regarding the content's complexity, the instructor's presence did not influence the cognitive load in the easy-topic videos (Wang et al., 2020a; Wang & Antonenko, 2017); however, the cognitive load was less in the instructor-presence videos with complex contents (Wang & Antonenko, 2017). Supporting this result, the intrinsic and extraneous cognitive load was lower in videos with the difficult topic but was not different for mental effort and germane load (Wang et al., 2020a). Additionally, videos including instructor with higher teaching enthusiasm (Qian et al., 2022) and professional coherence (Beege et al., 2022) decreased learners' cognitive load.

Regarding the model-observer similarity, the instructor's gender (Hoogerheide et al., 2018) and experience (Hoogerheide et al., 2016b) did not affect the cognitive load, but the age (Hoogerheide et al., 2016b) and body-orientation with a professional dress did (Beege et al., 2019). Additionally, male learners' cognitive load was less when they watched videos, including a male instructor (Hoogerheide et al., 2016a).

How the instructor is presented, content presentation style, whiteboard type, and the availability of nonverbal cues were other factors that might impact the cognitive load. Results were mixed for the instructor's presentation style. While Yi et al. (2019) showed that continuous presence negatively affected cognitive load, Kizilcec et al. (2015) reported that cognitive load was higher in strategic view than in constant view condition. Furthermore, the content's presentation style (i.e., drawing vs. static content) and the whiteboard type (i.e., transparent vs. conventional) (Fiorella et al., 2019), and nonverbal communication cues (i.e., gaze and gesture cues; Ouwehand et al., 2015) did not affect the cognitive load. However, the results for the beat and deictic gesture conditions were mixed (Beege et al., 2020).

### 4.1.3 Social learning outputs

**Social presence** Regarding the social presence, the following conditions were examined in the reviewed articles: video type, instructors' characteristics and presentation style, gaze behavior, and gestures. While Rosenthal and Walker (2020) demonstrated that the instructor's presence was influential on the learner's social presence (Study 2), some researchers (Alasfor, 2021; Colliot & Jamet, 2018; Ng & Przybyłek, 2021; Yuan et al., 2021) did not find a considerable effect. Additionally, the instructor's presence in the live composite videos was higher than in the picture-in-picture videos (Rosenthal & Walker, 2020, Study 2).

In two articles, the instructor's characteristics were compared. Learner's social presence was higher with the human instructor than with robot ones (Li et al., 2016). However, no significant difference was observed between the original and virtual instructor's views (Yuan et al., 2021).

Regarding the instructor's presence style in videos, no significant difference was found between continuous and intermitted presence (Yi et al., 2019). However, in one study (Kizilcec et al., 2015), strategic presence was superior to continuous presence. Finally, the instructor's gaze guidance, regardless of knowledge type (Wang et al., 2018) and deictic gestures (Beege et al., 2020) increased learners' social presence.

**Parasocial interaction** The gaze behavior, facial expressions (Pi et al., 2021), and instructor's proximity (Beege et al., 2017) did not have a meaningful effect on parasocial interaction. However, compared to the lateral orientation, the instructor's frontal orientation (Beege et al., 2017) and professional coherence (Beege et al., 2019, 2022) influenced learners' parasocial interaction. The deictic gesture was also effective on some dimensions of parasocial interaction (Beege et al., 2020).

## 4.2 Design-related issues of videos

The second research question focuses on the design-related issues of the videos in the reviewed articles (see Table 3). Regarding the video length, a total of 59 videos in 38 articles were considered since two articles (Alasfor, 2021; Hoogerheide et al., 2016a) did not report the video length, and one (Kizilcec et al., 2015) contained many videos with varying durations. Accordingly, the most frequently reported interval was 5 to 10 (42%), followed by 10 to 15 (29%), and lower than 5 min (24%).

As seen in Tables 3 and 65% of the reviewed articles focused on declarative knowledge. The percentage of procedural knowledge was 23, and 5% focused on both declarative and procedural knowledge.

The percentage of video styles were as follows picture-in-picture (52%), lecture-type (30%), and demonstration videos (4%). Additionally, 16% of videos was in voiceover format. In some articles, the instructor was presented strategically (3%), which means the instructor was not seen on the screen continuously. Lecture slides accompanied the instructor's image in most picture-in-picture videos (46%).

Table 3 also shows the contexts where the video lectures were captured. These videos were generally recorded in a lecture hall, followed by a setting where a board or a screen exists. Demonstration videos included a video model in front of a desk.

Although voiceover videos did not include the instructor's presence, they were used to compare them with other video styles. Lecture slides were generally used to present the content in these videos.

Instructors used different cues in the videos, including gaze (51%), gesture (37%), and facial expressions (16%). Among the gaze behaviors, directed gaze was the most used, followed by guided or shifted gaze. Regarding the gestures, instructors generally used pointing gestures. Facial expressions such as neutral, surprised, and happy faces were also used in the reviewed articles.

**Table 3** Design issues

Characteristic		n	%	
Video-related issues				
Video length				
	<= 4 min 59 s.	14	24	
	5 – 9 min 59 s.	25	42	
	10 – 14 min 59 s.	17	29	
	15 – 19 min 59 s.	1	2	
	>= 20 min	2	3	
Knowledge type				
	Declarative	28	65	
	Procedural	10	23	
	N/A	3	7	
	Declarative & procedural	2	5	
Video style				
Picture-in-picture	Instructor' continuous view	Lecture slide	74	46
		Khan-style	3	2
		Screencast	1	1
	Instructor's strategic view	4	3	
Video lecture	Lecture hall	11	7	
	Conventional whiteboard condition	9	6	
	Digital screen/curtain condition	7	4	
	Transparent whiteboard condition	5	3	
	Smartboard condition	4	3	
	Classroom	4	3	
	Empty room	4	3	
Voice-over	Seminar room	1	1	
	Lecture slide	19	12	
	Khan-style	2	1	
	Dynamic drawing on a board	2	1	
	Screencast	2	1	
Demonstration	Static drawing on the board	1	1	
	Video model in front of a desk	7	4	
Social cues				
Gaze	Direct gaze	10	25	
	Guided/shifted gaze	7	18	
	Averted gaze	2	5	
	Fixed gaze	1	3	
Gestures	Pointing gestures	5	13	
	Hand gestures	3	8	
	Beat gesture	3	8	
	Depictive gesture	3	8	
Face	General facial expressions	2	5	
	Neutral face	2	5	
	Surprised face	1	3	
	Happy face	1	3	

N/A refers to the number of articles that did not mention the design-related issue. *n* = number, % = percentage

## 5 Discussion

In this literature review, we examined the learning outcomes assessed in instructor-present videos and the design-related issues of these videos. We found that learners are more emotionally positive toward these videos. However, the results for cognitive and social aspects of learning are mixed. Additionally, design issues varied across video length, knowledge type, video style, and social or attentional cues the instructor used.

### 5.1 Affective, cognitive, and the social aspects of learning

Regarding the affective aspect of learning, learners' emotions, satisfaction, and learning enjoyment were relatively higher in instructor-presence videos. However, we observed that the video style was not influential on learners' emotions and satisfaction, and the instructor's characteristics, such as age, gender, and expertise, were not sufficiently effective on learning enjoyment. A possible explanation for the positive results may closely be related to the social cues provided by the instructor (Wang et al., 2020a), indicating that social cues have the potential to trigger the affective responses of learners. However, the instructor's age, gender, and expertise might not be equally valuable affective information or emotionally significant variables for learners in instructor-present videos. To test the assumption that higher similarity between the instructor and the learner enhances learners' affective states (Hoogerheide et al., 2016a), further research needs to address different instructor characteristics.

Attention, overall learning performance, retention, transfer, and cognitive load were the selected cognitive variables in the current study. Regarding attention, mixed findings were obtained for different video types. The instructor's face was a salient stimulus in the videos. Eye measurement data also showed that the instructor's gaze behavior could be used to draw learners' attention to the relevant learning element. These results seem consistent with other research that human faces attract a substantial amount of attention (Ouweland et al., 2015) and theoretical inferences indicating that the instructor's gaze behavior could be an attentional cue in multimedia learning settings (Mayer, 2005). Hence, it is plausible to assume that the instructor's presence may affect how students distribute their visual attention in instructional videos (Wang & Antonenko, 2017).

The results were inconclusive for learning performance, retention, and transfer in the reviewed articles. Many of the studies found instructor-present videos ineffective for learning-related variables. This finding is consistent with the results of previous works (Alemdag, 2022; Henderson & Schroeder, 2021) and the theoretical assumption of the image principle, stating that learning is not necessarily fostered by the instructor's image in multimedia learning environments (Mayer, 2014b). However, different variables such as learners' sustained attention (Kokoç et al., 2020), gestures (Beege et al., 2020), teaching enthusiasm (Qian et al., 2022), dynamic presentation on a board (Fiorella et al., 2019), instructor's professional congruence (Beege et al., 2022), presentation style (Yi et al., 2019), guided gaze behavior (Pi et al., 2020), contents' complexity (Wang et al., 2020a), and learner's prior knowledge (Fiorella

& Mayer, 2016) are the potential moderators for learning that need further examination in different conditions.

Contrary to findings reported by Alemdag (2022), we did not arrive at a definite conclusion about the learners' cognitive load in instructor-present videos. This inconsistency may be related to the measurement of cognitive load in the reviewed articles (Kizilcec et al., 2015). The other reason might be the design-related issues of instructor-present videos. From a cognitive load perspective, although the literature informs that the visual of the instructor creates extraneous processing (Wang & Antonenko, 2017), this study suggests that the learners' cognitive load can vary in instructor-present videos depending on the knowledge type (Hong et al., 2018), content's complexity (Wang et al., 2020a), and instructor's presentation style (Kizilcec et al., 2015; Yi et al., 2019).

The last learning dimension is the social learning outputs, which placed less attention in the reviewed articles than the affective and cognitive aspects of learning. Social presence was the most studied variable and was not relatively affected by the instructor's presence, as noticed by Alemdag (2022). An explanation for this result might be the instructor's voice, considered the most crucial social cue in learning environments (Colliot & Jamet, 2018). Alternatively, adding the instructor's image to the video may not truly enhance social presence (Yuan et al., 2021). Based on the reviewed articles' results, this study suggests that the instructor's presentation style (Kizilcec et al., 2015) and gaze guidance (Wang et al., 2018) may play a critical role in the learner's sense of social presence.

Regarding parasocial interaction, the reviewed articles reported inconclusive results for different conditions. Nevertheless, instructors' professional coherence was found more influential on parasocial processes, which shows the importance of professionalism in video lectures (Beege et al., 2019). To get more generalizable conclusions on learners' parasocial processes, more empirical research is required in which the instructor's presence is controlled in various contexts (Beege et al., 2022).

## 5.2 Design-related issues of videos

Each design issues of the instructor-present videos (see Table 3) might be a moderating variable affecting the learning outputs. This review indicated that the length of the instructor-present videos varied widely from 5 to 10 and 10 to 15 min. These lengths might not be a favorable learning condition for learners since the optimal video length for engagement is lower than six minutes for instructional videos (Guo et al., 2014). Alemdag (2022) also provided empirical evidence on this direction for instructor-present videos. In a long video, the instructor's dynamic visualization may increase the extraneous cognitive load and reduce learner engagement, leading to the consumption of working memory resources that need to be used for learning (Chen & Wu, 2015). This study, therefore, suggests that further empirical research is needed to determine the optimal video length conditions under which the instructor-present videos lead to better learning outcomes.

The knowledge type taught in the instructional videos is crucial for learning outcomes (Hong et al., 2018; Wang et al., 2020a). There may be changes in the

instructional roles of instructors while presenting different knowledge. In declarative knowledge presentations, for example, the instructor may only explain the relevant concepts verbally; however, in procedural knowledge presentations, it may be necessary to show the procedure step-by-step. Almost all reviewed articles in this study included videos containing either declarative or procedural knowledge presentations. A few articles compared both knowledge types concurrently, which might be an obstacle to drawing a more generalizable conclusion about the topic. Further research, therefore, is needed to understand how the instructor's presence affects learning outputs in different knowledge conditions.

The effect of the instructor's presence can vary depending on the video style (Chen & Wu, 2015; Wang & Antonenko, 2017). This review revealed that the most preferred video style was picture-in-picture, followed by lecture and demonstration videos. This finding may indicate a trend in using picture-in-picture videos in education (Li et al., 2016). In picture-in-picture videos, lecture slides were often combined with a talking head to convey learning content. Although its number is limited, the strategic presentation was also used in some picture-in-picture videos. In this presentation, the instructor is hidden when explaining critical contents (Yi et al., 2019). This study suggests this presentation style as an alternative solution to guide learners' attention and reduce the potential cognitive load during the instruction (Kizilcec et al., 2015).

Video lectures in the reviewed articles were usually created by recording the instructor giving a lecture in a lecture hall or a class where learners could observe the learner-learner and learner-instructor interactions. In these videos, the instructor explained the learning content by standing adjacent to the board. Because of that, these video lectures are considered to have a higher degree of media richness (Chen & Wu, 2015). The combined use of many media elements in these videos has raised the concern of a possible split-attention effect (Ouweland et al., 2015). However, it is possible to avoid such effect using social (e.g., guided gaze) or attentional cues (e.g., highlighting text). Therefore, the instructor can play a critical role in these videos.

In demonstration videos, the instructor usually stood behind a desk and solved a problem-solving task. Seeing the instructor's face in these videos was considered a distractive stimulus that may impede learning (van Wermeskerken & van Gog, 2017). However, the instructor's face might facilitate learning in a social interaction context where the instructor demonstrates the learning task (Van Gog et al., 2014). Hence, the instructor's non-verbal communication cues in these videos might have a higher effect on learning. However, relatively little is known about how to design demonstration videos to foster learning (Van Gog et al., 2014); therefore, further research on social communication cues is suggested.

The voiceover videos, which combine the instructor's narration with the learning content, were another frequently utilized video type in the reviewed articles. Examples of this type included the presentation of lecture slides, Khan-style, drawing on a board, and screencast. Researchers have often used this video to compare it with other instructor-present videos regarding various learning outcomes. The proponent of this format advocates that an instructor's presence may not be needed since it may cause an extraneous cognitive load hindering learning (Kizilcec et al., 2015; Wang

& Antonenko, 2017). However, learners do not have a positive attitude in this regard (Kizilcec et al., 2015; Wilson et al., 2018). Even some studies found instructor-present videos more effective than voiceover videos (Chen & Wu, 2015). It seems that the debate on the instructor's presence over different video styles will not end in the short term, and new research will continue to shape the field.

In recent years, there has been a considerable increase in studies focusing on social and attentional cues in instructor-present videos. This study categorized these cues into three groups: gaze behaviors, gestures, and facial expressions. From the social agency perspective, learners may not learn deeply from instructor-present videos (Mayer, 2014b). However, social cues may enhance learners' feeling of social presence (Mayer, 2014b) and motivate them to engage in generative processing for learning (Wang et al., 2020b), resulting in better learning performance. The literature suggests that as a social cue instructor's gaze behavior (Wang et al., 2018), gestures, and facial expressions (Stull et al., 2018) can play a critical role in learning. Additionally, the attentional cues may reduce the unnecessary cognitive load and help to consume working memory resources effectively (Ozcelik et al., 2010). Therefore, it is reasonable to conclude that both social and attentional cues should be used carefully in videos to establish a connection with the learners and direct their attention to the relevant part of the learning content. Moreover, when used properly, these cues may lay a bridge between the learner and the learning content over the instructor.

### 5.3 Practical and theoretical implications

The literature review yielded inconclusive results regarding the instructor's presence in videos. One of the reasons for these conflicting findings is likely the lack of detailed information describing the experimental conditions in the reviewed articles. In this regard, it becomes a requirement to collect more details about the instructors' characteristics, the learning context in which the video was prepared, the video style (e.g., demonstration or picture-in-picture videos), learners' characteristics who studied the videos, detail information about learning content, and the knowledge type being presented. Although some reviewed articles touched upon these factors, some neglected them. Therefore, the current analysis results might not be sufficient to make inferences for theory and practice related to the problem situation.

Nevertheless, the present study gives some insights into the design of instructor-present videos. Regarding the affective aspect of the learning, the instructor's presence might be a preferable design component in videos. Based on the results, it seems reasonable to suggest that instructor-present videos increase learners' positive feelings and satisfaction during the instruction. It is believed that this situation may have a favorable impact on learners' behaviors in online courses.

Regarding the cognitive aspect of learning, the instructor's presence may influence the learner's visual attention. However, the visual of the instructor may not contribute to the sense of social presence and the learning outputs, including learning performance, transfer, and retention. Nevertheless, using social and attentional cues in instructor-present videos may foster learning and trigger social responses

by learners. From a theoretical point of view, these cues might be suitable signaling components in video environments that lead learners to direct attention to the relevant learning content and help them allocate more cognitive resources to content. Furthermore, the results for learners' cognitive load were also inconclusive. In cases where the instructor is perceived as an additional information source, split attention or redundancy effects may occur. This study suggests a strategic presentation of the instructor to avoid these effects.

## 6 Conclusion

The main conclusion that can be drawn from this review is that design-related issues (i.e., video length, knowledge type, video style, non-verbal cues) of instructor-present videos may affect the different aspects of learning. Although the instructor's presence is favorable for affective learning outputs, the results for cognitive and social learning outputs are inconclusive and mixed. However, some social and attentional cues are suggested to benefit more from instructor-present videos in different learning conditions.

## Appendix 1: List of reviewed articles and the categories for synthesis

Authors (Year)	Learning outputs	Video Duration	Content of the video	Knowledge type	Context of the video	Cues
Alasfor (2021)	<ol style="list-style-type: none"> <li>1. Satisfaction</li> <li>2. Comprehension (Learning performance)</li> <li>3. Social presence</li> </ol>	N/A	Fundamentals of instructional technology	Declarative*	Video 1: Picture-in-picture presentation (PowerPoint+ talking head) Video 2: Voice-over presentation (in videoconferencing format)	N/A
Beege et al. (2017)	<ol style="list-style-type: none"> <li>1. Retention</li> <li>2. Transfer</li> <li>3. Parasocial interaction</li> </ol>	Three videos a. 2 min 38 s. b. 3 min 21 s. c. 3 min 45 s Total: 9 min 44 s	Statistics	Declarative*	Video 1: Lecturer (with frontal orientation and near proximity) in a lecture hall Video 2: Lecturer (with frontal orientation and far proximity) in a lecture hall Video 3: Lecturer (with lateral orientation and near proximity) in a lecture hall Video 4: Lecturer (with lateral orientation and far proximity) in a lecture hall	N/A

Authors (Year)	Learning outputs	Video Duration	Content of the video	Knowledge type	Context of the video	Cues
Beege et al. (2019)	1. Retention 2. Transfer 3. Cognitive load / Mental effort 4. Parasocial interaction	Experiment 1: 14 min 51 s. Experiment 2: 9 min 55 s.	Experiment 1: Multiple sclerosis Experiment 2: Amyotrophic lateral sclerosis	Experiment 1: Declarative* Experiment 2: Declarative*	Experiment 1: Video 1: Lecturer with a frontal professional appearance in a lecture hall Video 2: Lecturer with a lateral professional appearance in a lecture hall Video 3: Lecturer with a frontal non-professional appearance in a lecture hall Video 4: Lecturer with a lateral non-professional appearance in a lecture hall Experiment 2: Video 1: Lecturer with a frontal professional appearance in a lecture hall Video 2: Lecturer with a lateral professional appearance in a lecture hall Video 3: Lecturer with a frontal non-professional appearance in a lecture hall Video 4: Lecturer with a lateral non-professional appearance in a lecture hall	Experiment 1: N/A Experiment 2: N/A
Beege et al. (2020)	1. Emotion 2. Retention 3. Transfer 4. Cognitive load / Mental effort 5. Social presence 6. Parasocial interaction	Experiment 1: 9 min 14 s. Experiment 2: 9 min 20 s.	Experiment 1: Weather phenomena Experiment 2: Industrial revolution	Experiment 1: Declarative* Experiment 2: Declarative*	Experiment 1: Video 1: Picture-in-picture presentation (Instructor without gestures) Video 2: Picture-in-picture presentation (Instructor with beat gestures) Video 3: Picture-in-picture presentation (Instructor with deictic gestures) Experiment 2: Video 1: Picture-in-picture presentation (Instructor without gestures) Video 2: Picture-in-picture presentation (Instructor with beat gestures) Video 3: Picture-in-picture presentation (Instructor with deictic gestures)	Experiment 1: Beat gestures Deictic gestures Experiment 2: Beat gestures Deictic gestures

Authors (Year)	Learning outputs	Video Duration	Content of the video	Knowledge type	Context of the video	Cues
Beege et al. (2022)	1. Retention 2. Transfer 3. Cognitive load / Mental effort 4. Parasocial interaction	10 min 09 s.	Human nervous system	Declarative*	Video 1: Picture-in-picture presentation (Instructor with a professional appearance and communication) Video 2: Picture-in-picture presentation (Instructor with a professional appearance and non-professional-communication) Video 3: Picture-in-picture presentation (Instructor with a non-professional appearance and communication) Video 4: Picture-in-picture presentation (Instructor with a non-professional appearance and professional communication)	N/A
Chen and Wu (2015)	1. Emotion 2. Attention 3. Learning performance 4. Cognitive load / Mental effort	15 min.	Document Writing	Declarative*	Video 1: Lecturer in the class Video 2: Voice-over presentation (Powerpoint slide) Video 3: Picture-in-picture presentation (Powerpoint slide)	N/A
Colliot and Jamet (2018)	1. Attention 2. Retention 3. Transfer 4. Social presence	9 min 20 s.	Ebola (Biology)	Declarative*	Video 1: Picture-in-picture presentation Video 2: Voice-over presentation	N/A
Fiorella and Mayer (2016)	1. Retention 2. Transfer	1 min 40 s.	Dropper Effect (Physics)	Declarative*	Experiment 1: Video 1: Lecturer near a conventional whiteboard Video 2: Lecturer near a conventional whiteboard Video 3: Lecturer near a conventional whiteboard Experiment 4: Video 1: Lecturer near a conventional whiteboard Video 2: Lecturer near a conventional whiteboard	Hand moves Pointing

Authors (Year)	Learning outputs	Video Duration	Content of the video	Knowledge type	Context of the video	Cues
Fiorella et al. (2019)	1. Learning performance 2. Retention 3. Transfer 4. Cognitive load / Mental effort	Experiment 1: 12 min. Experiment 2: 12 min. Experiment 3: 12 min.	Study 1: Anatomy of Kidney (Biology) Study 2: Anatomy of Kidney (Biology) Study 3: Anatomy of Kidney (Biology)	Declarative*	Experiment 1: Video 1: Voice-over presentation (Static drawings on the board) Video 2: Voice-over presentation (Dynamic drawings -Khan style- on the board) Experiment 2: Video 1: Lecturer near a conventional whiteboard Video 2: Lecturer near a transparent whiteboard Experiment 3: Video 1: Voice-over presentation (Dynamic drawings -Khan style- on the board) Video 2: Lecturer near a transparent whiteboard	Experiment 1: N/A Experiment 2: Eye contact Experiment 3: Eye contact
Hew and Lo (2020)	1. Retention 2. Transfer	Study 1: 6 min. Study 2a: 5 min. Study 2b: 5 min.	Study 1: Arithmetic and Geometric Sequences (Math-Geometry) Study 2a: Rational and Irrational Numbers (Math-Geometry) Study 2b: Deductive Geometry (Geometry)	Procedural*	Study 1: Video 1: Lecturer in the class Video 2: Lecturer in a seminar room (PowerPoint + conversation video) Video 3: Voice-over presentation (PowerPoint + teacher's narration) Video 4: Voice-over presentation (PowerPoint + teacher's narration and photo) Video 5: Picture-in-picture presentation (PowerPoint + teacher's talking head) Video 6: Picture-in-picture presentation (Khan style + teacher's talking head) Study 2a: Video 1: Picture-in-picture presentation (PowerPoint + teacher's talking head) Video 2: Lecturer in the class Video 3: Picture-in-picture presentation (Khan style + teacher's talking head) Study 2b: Video 1: Picture-in-picture presentation (Khan style + teacher's talking head)	Study 1: N/A Study 2: N/A Study 3: N/A

Authors (Year)	Learning outputs	Video Duration	Content of the video	Knowledge type	Context of the video	Cues
Hong et al. (2018)	1. Learning performance 2. Cognitive load / Mental effort	Phase 1: 8 min 20 s. Phase 2: 7 min 60 s.	Phase 1: Educational technology Phase 2: Use of photoshop (computer science)	Phase 1: Declarative Phase 2: Procedural	Phase 1: Video 1: Voice-over presentation (Powerpoint slide) Video 2: Picture-in-picture presentation (Powerpoint slide) Phase 2: Video 1: Voice-over presentation (Screen recording) Video 2: Picture-in-picture presentation (Screen recording)	Phase 1: N/A Phase 2: N/A
Hoogerheide et al. (2018)	1. Learning enjoyment 2. Learning performance 3. Cognitive load / Mental effort	Type 1: 4 min Type 2: 4 min 4 s.	Electrical circuits (Physics)	Procedural*	Video 1: Lecturer (male) near to a screen (Powerpoint slide in the display) Video 2: Lecturer (male) near to a screen (Powerpoint slide in the display) Video 3: Lecturer (female) near to a screen (Powerpoint slide in the display) Video 4: Lecturer (female) near to a screen (Powerpoint slide in the display)	N/A
Hoogerheide et al. (2016a)	1. Learning enjoyment 2. Learning performance 3. Transfer 4. Cognitive load / Mental effort	N/A	Probability (Math)	Procedural*	Video 1: Video model (female) in front of a desk Video 2: Video model (male) in front of a desk	N/A
Hoogerheide et al. (2016b)	1. Learning enjoyment 2. Learning performance 3. Cognitive load / Mental effort	Type 1: 4 min Type 2: 4 min 4 s.	Electrical circuits (Physics)	Procedural*	Video 1: Lecturer (novice peer) near to a smartboard (Powerpoint slide in the display) Video 2: Lecturer (experienced peer) near to a smartboard (Powerpoint slide in the display) Video 3: Lecturer (novice adult) near to a smartboard (Powerpoint slide in the display) Video 4: Lecturer (experienced adult) near to a smartboard (Powerpoint slide in the display)	N/A

Authors (Year)	Learning outputs	Video Duration	Content of the video	Knowledge type	Context of the video	Cues
Kizilcec et al. (2015)	1. Emotion 2. Learning performance 3. Cognitive load / Mental effort 4. Social presence	Not specially presented. Phase 1: Six video lectures between 8 to 20 min. per week for eight weeks period. Phase 2: Six video lectures between 8 to 20 min. per week	Phase 1: Sociology Phase 2: Sociology	Declarative*	Phase 1: Video 1: Voice-over presentation (Powerpoint slide) Video 2: Picture-in-picture presentation (Powerpoint slide) Phase 2: Video 1: Picture-in-picture presentation (Static view of the instructor on the screen) Video 2: Picture-in-picture presentation (Strategic view of the instructor on the screen)	Phase 1: N/A Phase 2: N/A
Kokoç et al. (2020)	1. Learning performance	3 min.	Sequencing in Python	N/A	Video 1: Picture-in-picture presentation (ppt slide + talking head) Video 2: Voice-over presentation Video 3: Screencast lecture	N/A
Korving et al. (2016)	1. Attention	Two video lectures. Approx. 7 min 30 s for each video.	N/A	N/A	Video 1: Picture-in-picture presentation (Maximal visibility of the instructor [in a lecture hall] on the screen) Video 2: Picture-in-picture presentation (Minimal visibility of the instructor [in a lecture hall] on the screen) Video 3: Voice-over presentation	N/A
Lackmann et al. (2021)	1. Attention 2. Learning performance	14 min.	Introduction to organizational behavior	Declarative*	Video 1: Voice-over presentation (PowerPoint slide*) Video 2: Lecturer in a classroom	N/A
Li et al. (2016)	1. Emotion 2. Retention 3. Social presence	Approx. 13 min.	Human-computer interaction (computer science)	N/A	Video 1: Picture-in-picture presentation (with human model) Video 2: Picture-in-picture presentation (with an animated human) Video 3: Picture-in-picture presentation (with a robot) Video 4: Picture-in-picture presentation (with an animated robot)	N/A
Ng and Przybylek, (2021)	1. Retention 2. Cognitive load / Mental effort 3. Social presence	10 min.	learning Mandarin Chinese vocabularies	Declarative*	Video 1: Voice-over presentation (PowerPoint slide) Video 2: Picture-in-picture presentation (PowerPoint slide)	N/A

Authors (Year)	Learning outputs	Video Duration	Content of the video	Knowledge type	Context of the video	Cues
Ouwehand et al. (2015)	1. Attention 2. Learning performance 3. Transfer 4. Cognitive load / Mental effort	2 min.	Problem-solving task (not including any specific content: three-step water-redistribution problem)	Procedural*	Video 1: Picture-in-picture presentation (Powerpoint slide + no cue) Video 2: Picture-in-picture presentation (Powerpoint slide + gaze cue) Video 3: Picture-in-picture presentation (Powerpoint slide + gesture and gaze cue)	Gaze guidance Pointing gesture
Pi et al. (2019)	1. Attention 2. Retention 3. Transfer	8 min 40 s.	Cloning (Biology)	Declarative*	Video 1: Picture-in-picture presentation (Powerpoint slide) Video 2: Picture-in-picture presentation (Powerpoint slide) Video 3: Picture-in-picture presentation (Powerpoint slide) Video 4: Picture-in-picture presentation (Powerpoint slide)	Direct gaze Pointing gesture by hand
Pi et al. (2020)	1. Attention 2. Retention 3. Transfer	7 min 20 s.	A Geography content (wind)	Declarative*	Video 1: Picture-in-picture presentation (Powerpoint slide) Video 2: Picture-in-picture presentation (Powerpoint slide) Video 3: Picture-in-picture presentation (Powerpoint slide) Video 4: Picture-in-picture presentation (Powerpoint slide) Video 5: Picture-in-picture presentation (Powerpoint slide) Video 6: Picture-in-picture presentation (Powerpoint slide)	Direct gaze Guided gaze Averted gaze
Pi et al. (2021)	1. Attention 2. Learning performance 3. Parasocial interaction	8 min.	Earthquakes	Declarative*	Video 1: Picture-in-picture presentation (instructor with direct gaze and neutral face + ppt slide) Video 2: Picture-in-picture presentation (instructor with direct gaze and surprised face + ppt slide) Video 3: Picture-in-picture presentation (instructor with guided gaze and neutral face + ppt slide) Video 4: Picture-in-picture presentation (instructor with guided gaze and surprised face + ppt slide)	Direct gaze Gaze guidance Neutral face Surprised face

Authors (Year)	Learning outputs	Video Duration	Content of the video	Knowledge type	Context of the video	Cues
Pi et al. (2022a)	1. Attention 2. Learning performance	7 min.	Earth	Declarative*	<p>Video 1: Picture-in-picture presentation (instructor with direct gaze and happy face + ppt slide)</p> <p>Video 2: Picture-in-picture presentation (instructor with direct gaze and neutral face + ppt slide)</p> <p>Video 3: Picture-in-picture presentation (instructor with averted gaze and happy face + ppt slide)</p> <p>Video 4: Picture-in-picture presentation (instructor with averted gaze and neutral face + ppt slide)</p>	<p>Direct gaze</p> <p>Averted gaze</p> <p>Happy face</p> <p>Neutral face</p>

Authors (Year)	Learning outputs	Video Duration	Content of the video	Knowledge type	Context of the video	Cues
Pi et al. (2022b)	1. Learning performance	Experiment 1: 4 min. Experiment 2: 13 min.	Experiment 1: English learning Experiment 2: English learning	Experiment 1: Declarative* Experiment 2: Declarative*	Experiment 1: Video 1: Picture-in-picture presentation (instructor with beat gesture + ppt slide) Video 2: Picture-in-picture presentation (instructor with pointing gesture + ppt slide) Video 3: Picture-in-picture presentation (instructor with depictive gesture + ppt slide) Experiment 2: Video 1: Picture-in-picture presentation (instructor with pointing gesture + ppt slide with simple visual material) Video 2: Picture-in-picture presentation (instructor with pointing gesture + ppt slide with complex visual material) Video 3: Picture-in-picture presentation (instructor without pointing gesture + ppt slide with simple visual material) Video 4: Picture-in-picture presentation (instructor without pointing gesture + ppt slide with complex visual material)	Experiment 1: Beat gesture Pointing gesture Depictive gesture Experiment 2: Pointing gesture
Qian et al. (2022)	1. Attention 2. Learning performance 3. Cognitive load / Mental effort	13 min 25 s.	A chemistry content	Declarative	Video 1: Lecturer with a high teaching enthusiasm near a digital curtain Video 1: Lecturer with a low teaching enthusiasm near a digital curtain	N/A

Authors (Year)	Learning outputs	Video Duration	Content of the video	Knowledge type	Context of the video	Cues
Rosenthal and Walker (2020)	1. Emotion 2. Attention 3. Retention 4. Social presence	Study1: 6 min 17 s. – 6 min 25 s. 6 min 46 s. – 7 min 09 s. Study2: 12 min 31 s. & 14 min 11 s. 12 min 14 s. & 13 min 16 s. 14 min 4 s. & 15 min 36 s.	Study 1: 1. Effective learning (education) 2. Correlation and causation (statistics) Study 2: 1. environmental sustainability	Declarative*	Study 1: Video 1: Lecturer near to a screen (Powerpoint slide in the display) Video 2: Picture-in-picture presentation (Powerpoint slide) Video 3: Voice-over presentation Video 4: Picture-in-picture presentation (with live composite style) Study 2: Video 1: Picture-in-picture presentation (Powerpoint slide) Video 2: Voice-over presentation Video 3: Picture-in-picture presentation (with live composite style)	Study 1: N/A Study 2: N/A
Stull et al. (2018)	1. Attention 2. Learning performance	20 min.	Organic chemistry	Declarative*	Video 1: Lecturer near a conventional whiteboard Video 2: Lecturer near a transparent whiteboard	Direct gaze Gaze guidance
Stull et al. (2021)	1. Attention	14 min.	Kidney anatomy and physiology (biology)	Declarative*	Video 1: Lecturer (with shifted gaze behavior) near to a transparent whiteboard Video 2: Lecturer (with fixed gaze behavior) near to a transparent whiteboard Video 3: Lecturer (with shifted gaze behavior) near to a conventional whiteboard Video 4: Lecturer (with fixed gaze behavior) near to a conventional whiteboard	Shifting gaze Fixed gaze
Van Gog et al. (2014)	1. Attention 2. Learning performance 3. Transfer	2 min 22 s.	Problem-solving task (not including any specific content)	Procedural*	Video 1: Video model (with her face) in front of a desk Video 2: Video model (without her face) in front of a desk	N/A
van Wermeskerken and van Gog (2017)	1. Attention 2. Learning performance	4 min 7 s.	Chemistry	Procedural*	Video 1: Video model (without face) in front of a desk Video 2: Video model (with face and gaze guidance) in front of a desk Video 3: Video model (with face and without gaze guidance) in front of a desk	Gaze guidance

Authors (Year)	Learning outputs	Video Duration	Content of the video	Knowledge type	Context of the video	Cues
van Wermeskerken et al. (2018)	1. Attention 2. Retention 3. Transfer	8 min 57 s.	Probability (Math)	Procedural*	Video 1: Picture-in-picture presentation (Powerpoint slide) Video 2: Voice-over presentation	Direct gaze
Wang and Antonenko (2017)	1. Satisfaction 2. Attention 3. Retention 4. Transfer 5. Cognitive load / Mental effort	10 min.	Geometry	Procedural*	Video 1: Picture-in-picture presentation for an easy topic Video 2: Voice-over presentation for an easy topic (Khan style) Video 3: Picture-in-picture presentation for a difficult topic Video 4: Voice-over presentation for a difficult topic (Khan style)	N/A
Wang et al. (2018)	1. Attention 2. Learning performance 3. Social presence	Two videos: For declarative knowledge 6 min 55 s. For procedural knowledge 6 min 54 s.	1. Red, green, and blue color model (computer science) 2. Image composition (computer science)	Video 1: Declarative Video 2: Procedural	Video 1: Picture-in-picture presentation (for declarative knowledge + without gaze) Video 2: Picture-in-picture presentation (for declarative knowledge + with gaze guidance) Video 3: Picture-in-picture presentation (for procedural knowledge + without gaze) Video 4: Picture-in-picture presentation (for procedural knowledge + with gaze guidance)	Guided gaze Direct gaze
Wang et al. (2020)	1. Emotion 2. Satisfaction 3. Attention 4. Retention 5. Transfer 6. Cognitive load / Mental effort	Two videos Easy topic 3 min. Difficult topic 4 min.	Statistics	Declarative*	Video 1: Picture-in-picture presentation (for an easy topic) Video 2: Voice-over presentation (for an easy topic) Video 3: Picture-in-picture presentation (for a difficult topic) Video 4: Voice-over presentation (for a difficult easy topic)	eye-gaze gesture facial expression
Wang et al. (2020b)	1. Retention 2. Transfer 3. Cognitive load / Mental effort	Two videos. Each one is 4 min.	Statistics	Declarative*	Video 1: Picture-in-picture presentation (Powerpoint slide) Video 2: Voice-over presentation (Powerpoint slide)	**Facial expression **Body gesturing **Mutual eye contact

Authors (Year)	Learning outputs	Video Duration	Content of the video	Knowledge type	Context of the video	Cues
Wilson et al. (2018)	1. Learning enjoyment 2. Comprehension (Learning performance)	Study 1: 24 min Study 2: 6 min 30 s. Study 3: 6 min 30 s Study 4: 6 min 30 s	Study 1: Biology (global problems of population growth) Study 2: History Study 3: History Study 4: History	Declarative*	Study 1: Video 1: Audio lecture Video 2: Audio lecture with text Video 3: Lecturer in a lecture hall Study 2: Video 1: Audio lecture Video 2: Lecturer in a lecture hall Study 3: Video 1: Audio lecture Video 2: Text lecture Video 3: Audio lecture with text Video 3: Lecturer in a lecture hall Study 4: Video 1: Audio lecture Video 2: Text lecture Video 3: Audio lecture with text Video 3: Picture-in-picture (Lecturer in a lecture hall + text presentation)	Experiment 1: N/A Experiment 2: N/A Experiment 3: N/A Experiment 4: N/A
Yi et al. (2019)	1. Satisfaction 2. Retention 3. Transfer 4. Cognitive load / Mental effort 5. Social presence	4 min 17 s.	Educational technology (Preparing a video in Camtasia)	Procedural	Video 1: Picture-in-picture presentation (Continuous view of the instructor on the screen) Video 2: Picture-in-picture presentation (Intermittent view of the instructor on the screen)	N/A
Yuan et al. (2021)	1. Emotion 2. Satisfaction 3. Retention 4. Transfer 5. Social presence	9 min	Describing the meaning, origin, and related allusions of paleogeography	Declarative*	Video 1: Picture-in-picture presentation (PowerPoint Slide + original view of instructor) Video 2: Picture-in-picture presentation (PowerPoint Slide + face-beautified view of instructor) Video 3: Picture-in-picture presentation (PowerPoint Slide + virtual instructor view) Video 4: Voice-over presentation (PowerPoint Slide)	N/A

Authors (Year)	Learning outputs	Video Duration	Content of the video	Knowledge type	Context of the video	Cues
Zhang et al. (2021)	1. Satisfaction 2. Attention 3. Learning performance	5 min.	Topic of 'Sleep' in Chinese	Declarative*	Video 1: Voice-over presentation Video 2: Picture-in-picture presentation (The instructor is on the left) Video 3: Picture-in-picture presentation (The instructor is in the middle) Video 4: Video 2: Picture-in-picture presentation (The instructor is on the right)	N/A

\* Although these communication cues were used by the lecturer, the detail of how authors used them was not presented

\*\* Although these communication cues were used by the lecturer, the detail of how authors used them was not presented

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1007/s10639-022-11532-4>.

**Data availability** The datasets generated during and/or analyzed during the current study has associated in a repository in the Appendix 1.

## Declarations

**Conflict of interest** The author has no relevant financial or non-financial interests to disclose.

## References

- Alasfor, K. (2021). Impact of showing a male instructor's face on female students' perceived social presence, satisfaction, and comprehension in distance education in a conservative, gender-segregated culture. *Educational Technology Research and Development*, 69(3), 1799–1810. <https://doi.org/10.1007/s11423-021-10013-8>
- Alemdag, E. (2022). Effects of instructor-present videos on learning, cognitive load, motivation, and social presence: a meta-analysis. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-022-11154-w>
- Beege, M., Kriegelstein, F., & Arnold, C. (2022). How instructors influence learning with instructional videos - The importance of professional appearance and communication. *Computers and Education*, 185. <https://doi.org/10.1016/j.compedu.2022.104531>
- Beege, M., Nebel, S., Schneider, S., & Rey, G. D. (2019). Social entities in educational videos: combining the effects of addressing and professionalism. *Computers in Human Behavior*, 93, 40–52. <https://doi.org/10.1016/j.chb.2018.11.051>
- Beege, M., Ninaus, M., Schneider, S., Nebel, S., Schlemmel, J., Weidenmüller, J., Moeller, K., & Rey, G. D. (2020). Investigating the effects of beat and deictic gestures of a lecturer in educational videos. *Computers and Education*, 156. <https://doi.org/10.1016/j.compedu.2020.103955>
- Beege, M., Schneider, S., Nebel, S., & Rey, G. D. (2017). Look into my eyes! Exploring the effect of addressing in educational videos. *Learning and Instruction*, 49, 113–120. <https://doi.org/10.1016/j.learninstruc.2017.01.004>

- Chen, C. M., & Wu, C. H. (2015). Effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance. *Computers and Education*, *80*, 108–121. <https://doi.org/10.1016/j.compedu.2014.08.015>
- Colliot, T., & Jamet, É. (2018). Understanding the effects of a teacher video on learning from a multimedia document: an eye-tracking study. *Educational Technology Research and Development*, *66*(6), 1415–1433. <https://doi.org/10.1007/s11423-018-9594-x>
- Denyer, D., & Tranfield, D. (2009). Producing a systematic review. In D. A. Buchanan, & A. Bryman (Eds.), *The sage handbook of organizational research methods* (pp. 671–689). Sage Publications Ltd.
- Fiorella, L., & Mayer, R. E. (2016). Effects of observing the instructor draw diagrams on learning from multimedia messages. *Journal of Educational Psychology*, *108*(4), 528–546. <https://doi.org/10.1037/edu0000065>
- Fiorella, L., Stull, A. T., Kuhlmann, S., & Mayer, R. E. (2019). Instructor presence in video lectures: the role of dynamic drawings, eye contact, and instructor visibility. *Journal of Educational Psychology*, *111*(7), 1162–1171. <https://doi.org/10.1037/edu0000325>
- Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: An empirical study of MOOC videos. *Proceedings of the First ACM Conference on Learning @ Scale Conference*, 41–50. <https://doi.org/10.1145/2556325.2566239>
- Henderson, M. L., & Schroeder, N. L. (2021). A systematic review of instructor presence in instructional videos: Effects on learning and affect. *Computers and Education Open*, *2*(October), 100059. <https://doi.org/10.1016/j.caeo.2021.100059>
- Hew, K. F., & Lo, C. K. (2020). Comparing video styles and study strategies during video-recorded lectures: effects on secondary school mathematics students' preference and learning. *Interactive Learning Environments*, *28*(7), 847–864. <https://doi.org/10.1080/10494820.2018.1545671>
- Hong, J., Pi, Z., & Yang, J. (2018). Learning declarative and procedural knowledge via video lectures: cognitive load and learning effectiveness. *Innovations in Education & Teaching International*, *55*(1), 74–81.
- Hoogerheide, V., Loyens, S. M. M., & van Gog, T. (2016a). Learning from video modeling examples: does gender matter? *Instructional Science*, *44*(1), 69. <https://doi.org/10.1007/s11251-015-9360-y>
- Hoogerheide, V., van Wermeskerken, M., Loyens, S. M. M., & van Gog, T. (2016b). Learning from video modeling examples: Content kept equal, adults are more effective models than peers. *Learning & Instruction*, *44*, 22–30. <https://doi.org/10.1016/j.learninstruc.2016.02.004>
- Hoogerheide, V., van Wermeskerken, M., van Nassau, H., & van Gog, T. (2018). Model-observer similarity and task-appropriateness in learning from video modeling examples: do model and student gender affect test performance, self-efficacy, and perceived competence? *Computers in Human Behavior*, *89*, 457–464. <https://doi.org/10.1016/j.chb.2017.11.012>
- Kizilcec, R. F., Bailenson, J. N., & Gomez, C. J. (2015). The instructor's face in video instruction: evidence from two large-scale field studies. *Journal of Educational Psychology*, *107*(3), 724–739. <https://doi.org/10.1037/edu0000013>
- Kizilcec, R. F., Papadopoulos, K., & Sritanyaratana, L. (2014). Showing face in video instruction: Effects on information retention, visual attention, and affect. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2095–2102. <https://doi.org/10.1145/2556288.2557207>
- Kokoç, M., Ilgaz, H., Altun, A., Kokoc, M., Ilgaz, H., Altun, A., Kokoç, M., Ilgaz, H., & Altun, A. (2020). Effects of sustained attention and video lecture types on learning performances. *Educational Technology Research and Development*, *68*(6), 3015–3039. <https://doi.org/10.1007/s11423-020-09829-7>
- Korving, H., Hernández, M., & De Groot, E. (2016). Look at me and pay attention! A study on the relation between visibility and attention in weblectures. *Computers and Education*, *94*, 151–161. <https://doi.org/10.1016/j.compedu.2015.11.011>
- Lackmann, S., Léger, P. M., Charland, P., Aubé, C., & Talbot, J. (2021). The influence of video format on engagement and performance in online learning. *Brain Sciences*, *11*(2), 1–21. <https://doi.org/10.3390/brainsci11020128>
- Li, J., Kizilcec, R., Bailenson, J., & Ju, W. (2016). Social robots and virtual agents as lecturers for video instruction. *Computers in Human Behavior*, *55*, 1222–1230. <https://doi.org/10.1016/j.chb.2015.04.005>
- Linnenluecke, M. K., Marrone, M., & Singh, A. K. (2019). Conducting systematic literature reviews and bibliometric analyses. *Australian Journal of Management*, *45*(2), 175–194. <https://doi.org/10.1177/0312896219877678>

- Mayer, R. E. (2005). Cognitive theory of Multimedia Learning. *The Cambridge handbook of multimedia learning* (pp. 31–48). Cambridge University Press. <https://doi.org/10.1017/CBO9780511816819.004>.
- Mayer, R. E. (2014a). Multimedia instruction. In J. M. Spector (Ed.), *Handbook of research on educational communications and technology* (pp. 385–399). Springer.
- Mayer, R. E. (2014b). Principles of multimedia learning based on social cues: personalization, voice, and image principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 419–447). Cambridge University Press.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, T. P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLOS Medicine*, 6(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Moon, J., & Ryu, J. (2020). The effects of social and cognitive cues on learning comprehension, eye-gaze pattern, and cognitive load in video instruction. *Journal of Computing in Higher Education*, 33, 39–63. <https://doi.org/10.1007/s12528-020-09255-x>
- Ng, Y. Y., & Przybylek, A. (2021). Instructor presence in video lectures: preliminary findings from an online experiment. *IEEE Access: Practical Innovations, Open Solutions*, 9, 36485–36499. <https://doi.org/10.1109/ACCESS.2021.3058735>
- Ouweland, K., van Gog, T., & Paas, F. (2015). Designing effective video-based modeling examples using gaze and gesture cues. *Educational Technology & Society*, 18(4), 78–88.
- Ozcelik, E., Arslan-Ari, I., & Cagiltay, K. (2010). Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior*, 26(1), 110–117. <https://doi.org/10.1016/j.chb.2009.09.001>
- Pi, Z., Chen, M., Zhu, F., Yang, J., & Hu, W. (2022a). Modulation of instructor's eye gaze by facial expression in video lectures. *Innovations in Education and Teaching International*, 59(1), 15–23. <https://doi.org/10.1080/14703297.2020.1788410>
- Pi, Z., Zhang, Y., Yu, Q., Zhang, Y., Yang, J., & Zhao, Q. (2022b). Neural oscillations and learning performance vary with an instructor's gestures and visual materials in video lectures. *British Journal of Educational Technology*, 53(1), 93–113. <https://doi.org/10.1111/bjet.13154>
- Pi, Z., & Hong, J. (2016). Learning process and learning outcomes of video podcasts including the instructor and PPT slides: a chinese case. *Innovations in Education and Teaching International*, 53(2), 135–144. <https://doi.org/10.1080/14703297.2015.1060133>
- Pi, Z., Xu, K., Liu, C., & Yang, J. (2020). Instructor presence in video lectures: Eye gaze matters, but not body orientation. *Computers & Education*, 144(152), <https://doi.org/10.1016/j.compedu.2019.103713>.
- Pi, Z., Zhang, Y., Zhu, F., Chen, L., Guo, X., & Yang, J. (2021). The mutual influence of an instructor's eye gaze and facial expression in video lectures. *Interactive Learning Environments*. <https://doi.org/10.1080/10494820.2021.1940213>
- Pi, Z., Zhang, Y., Zhu, F., Xu, K., Yang, J., & Hu, W. (2019). Instructors' pointing gestures improve learning regardless of their use of directed gaze in video lectures. *Computers and Education*, 128(October 2018), 345–352. <https://doi.org/10.1016/j.compedu.2018.10.006>
- Qian, H., Li, H., Tang, S., Wang, J., Liu, Q., & Chen, G. (2022). How teacher enthusiasm affects students' learning of chemistry declarative knowledge in video lectures. *Chemistry Education Research and Practice*, 1. <https://doi.org/10.1039/d2rp00095d>
- Rosenthal, S., & Walker, Z. (2020). Experiencing live composite video lectures: comparisons with traditional lectures and common video lecture methods. *International Journal for the Scholarship of Teaching & Learning*, 14(1), 1–10.
- Schneider, S., Beege, M., Nebel, S., Schnaubert, L., & Rey, G. D. (2022). The cognitive-affective-social theory of learning in digital environments (CASTLE). *Educational Psychology Review*, 34(1), 1–38. <https://doi.org/10.1007/s10648-021-09626-5>
- Stull, A. T., Fiorella, L., & Mayer, R. E. (2018). An eye-tracking analysis of instructor presence in video lectures. *Computers in Human Behavior*, 88(June), 263–272. <https://doi.org/10.1016/j.chb.2018.07.019>
- Stull, A. T., Fiorella, L., & Mayer, R. E. (2021). The case for embodied instruction: the instructor as a source of attentional and social cues in video lectures. *Journal of Educational Psychology*, 113(7), 1441–1453. <https://doi.org/10.1037/edu0000650>
- Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management*, 14(3), 207–222. <https://doi.org/10.1111/1467-8551.00375>

- Van Gog, T., Verveer, I., & Verveer, L. (2014). Learning from video modeling examples: Effects of seeing the human model's face. *Computers and Education*, 72, 323–327. <https://doi.org/10.1016/j.compedu.2013.12.004>
- van Wermeskerken, M., Ravensbergen, S., & van Gog, T. (2018). Effects of instructor presence in video modeling examples on attention and learning. *Computers in Human Behavior*, 89, 430–438. <https://doi.org/10.1016/j.chb.2017.11.038>
- van Wermeskerken, M., & van Gog, T. (2017). Seeing the instructor's face and gaze in demonstration video examples affects attention allocation but not learning. *Computers and Education*, 113, 98–107. <https://doi.org/10.1016/j.compedu.2017.05.013>
- Wang, H., Pi, Z., & Hu, W. (2018). The instructor's gaze guidance in video lectures improves learning. *Journal of Computer Assisted Learning*. <https://doi.org/10.1111/jcal.12309>
- Wang, J., & Antonenko, P. D. (2017). Instructor presence in instructional video: Effects on visual attention, recall, and perceived learning. *Computers in Human Behavior*, 71, 79–89. <https://doi.org/10.1016/j.chb.2017.01.049>
- Wang, J., Antonenko, P., & Dawson, K. (2020a). Does visual attention to the instructor in online video affect learning and learner perceptions? An eye-tracking analysis. *Computers & Education*, 146. <https://doi.org/10.1016/j.compedu.2019.103779>
- Wang, J., Antonenko, P., Keil, A., & Dawson, K. (2020b). Converging subjective and psychophysiological measures of cognitive load to study the effects of instructor-present video. *Mind Brain and Education*, 14(3), 279–291. <https://doi.org/10.1111/mbe.12239>
- Wilson, K. E., Martinez, M., Mills, C., D'Mello, S., Smilek, D., & Risko, E. F. (2018). Instructor presence effect: liking does not always lead to learning. *Computers & Education*, 122(March), 205–220. <https://doi.org/10.1016/j.compedu.2018.03.011>
- Yi, T., Yang, X., Pi, Z., Huang, L., & Yang, J. (2019). Teachers' continuous vs. intermittent presence in procedural knowledge instructional videos. *Innovations in Education and Teaching International*, 56(4), 481–492. <https://doi.org/10.1080/14703297.2018.1470020>
- Yuan, M., Zeng, J., Wang, A., & Shang, J. (2021). Would it be better if instructors technically adjust their image or voice in online courses? Impact of the way of instructor presence on online learning. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.746857>
- Zhang, Y., Xu, K., Pi, Z., & Yang, J. (2021). Instructor's position affects learning from video lectures in chinese context: an eye-tracking study. *Behaviour and Information Technology*, 1–10. <https://doi.org/10.1080/0144929X.2021.1910731>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.