Alemdag *Research and Practice in Technology Enhanced Learning* (2023) 18:29
https://doi.org/10.58459/rptel.2023.18029

Research and Practice in
Technology Enhanced Learning

RESEARCH **Free and Open Access**

A scoping review of the literature on embodied instructional videos

Ecenaz Alemdag

ecenazalemdag@gmail.com
Department of Computer Education and Instructional Technology, Middle East Technical University, Ankara, Turkey

Abstract

Although it is prevalent to use embodied modes of information (e.g., instructors' movements and gestures) in instructional videos, there is a lack of comprehensive review elucidating how this type of information is designed and investigated in research studies. This scoping review of the literature examined 55 empirical research articles with 71 separate studies regarding embodied instructional videos to reveal their characteristics and design factors and provide key findings regarding their effects. The results revealed that most videos included slides for the lecture and demonstration of science subjects. The following design factors were determined in embodied instructional videos: instructor demeanor, instructor visual presence, generative activities, learner characteristics, content, instructional media, and scene. The findings regarding their influence on learning supported existing embodiment principles. They also uncovered the contributing or moderating effect of instructors' deictic gestures, facial expressions, and intimate behaviors, students' prior knowledge and actions during learning, and the complexity of the video subject. Overall, this review provides helpful information for practitioners based on empirical evidence and indicates research gaps in the literature on embodied instructional videos.

Keywords: Educational video, Embodied instruction, Embodiment, Instructional video, Scoping review

Introduction

Instructional videos are multimedia learning materials with both visual (e.g., illustrations, graphs, and animations) and auditory (e.g., sound effects and instructors' voice) information that can enhance the processing and learning of content (Lange & Costly, 2020). They have become quite widespread in classroom-based, blended, distance learning, and massive open online courses (MOOCs) (Belt & Lowenthal, 2021; Harrison, 2020; Pi et al., 2021d; Polat et al., 2021). The low expense of video authoring, hosting, and streaming technologies and the accessibility of videos anywhere and anytime led instructors to prepare and use instructional videos, especially in higher education (Fyfield et al., 2019; Ou et al., 2019). However, students often encounter trouble sustaining their attention and understanding the content of instructional videos when their content and delivery are inefficiently organized (Tseng, 2021). Therefore, instructors need to be recognizant of potentially problematic issues regarding the design of instructional videos and eliminate them before using videos in their lessons (Lange & Costly, 2020).

Embodiment is one aspect of instructional video design that needs to be considered by practitioners. It is concerned with instructors' movements, gestures, and uses of instructional tools and classroom space (Lim, 2021). Its function is to provide students with social and attentional cues so that they can feel a social partnership with the instructor and become oriented to follow visual content while watching instructional videos (Stull et al., 2021). According to the embodiment principle, people learn more deeply when such cues are integrated into multimedia learning environments (Mayer, 2014b). Moreover, social-agency theory suggests that learners can feel more motivated for deep processing of verbal and visual modes of information with these cues because they can perceive their interaction with the computer as human-to-human conversation and develop a partnership with the computer (Atkinson et al., 2005; Mayer et al., 2003). Furthermore, instructors' embodiment in videos including their tone of voice, body language, and facial expression were found to be more important for students than the technical quality of videos (Harrison, 2020). Hence, the embodiment is a critical design element in instructional video development to enhance information processing, learning, and motivation (Pi et al., 2021d).

There is also an increasing trend in the number of studies examining embodiment in instructional videos (Pi et al., 2021e; Schneider et al., 2021). Mayer (2021) synthesized the research conducted by him and his colleagues and proposed four suggestions collected under the embodiment principle. They supported instructors' dynamic drawing, gaze guidance, dynamic gestures, and the first-person perspective to improve learning from videos. However, this synthesis has remained limited because it did not cover the whole spectrum of existing research on embodied instructional videos. Moreover, Henderson and Schroeder (2021) performed a systematic review study to examine the effect of on-screen instructors in videos. They indicated contradictory research findings regarding whether instructors should be present or absent in videos to enhance learning and decrease cognitive load. They called for future review studies to deeply investigate the embodiment principle and determine certain instructor behaviors and appearances beneficial for learning. Therefore, a comprehensive synthesis of the studies on embodied instructional videos can be helpful to fill the gaps in the literature. Accordingly, this scoping review of the literature aims to analyze empirical studies regarding embodied instructional videos in terms of demographics, video characteristics, design factors, and outcome variables and indicate how the design of these videos influences the learning process and outcomes.

Background

Embodiment in multimedia learning

Multimodal or multimedia learning occurs when learners construct knowledge from words and pictures (Moreno & Mayer, 2007). However, aids for meaning making are not limited to words and pictures in educational environments. They can involve instructors' movements, gestures, and use of instructional tools and classroom space, especially in teaching abstract subjects such as math (Alibali & Nathan, 2012). These nonverbal communication resources are embodied modes in education or educational semiotics (Lim, 2021). It is necessary to consider embodied modes with their potentials and limitations and how they can be integrated with words and pictures to benefit from multimodality or utilization of different modes for the construction of meaning in learning (Jewitt et al., 2016).

In the multimedia learning field, social-agency and embodied instruction theories explain the learning mechanism with embodied modes. According to these theories, instructors are considered social agents who provide social and attentional cues to learners with embodied modes such as gestures, movements, eye contact, and facial expressions (Stull et al., 2021). Social cues such as direct gaze help learners adopt the instructor as a social partner making conversations with them (Mayer & DaPra, 2012). More particularly, social agency theory suggests that learners will perceive the instructor with social cues in the multimedia learning environment as a person giving an effort for learners to make sense of information (Atkinson et al., 2005). As a result, learners can feel more motivated to select, organize, and integrate relevant verbal and visual modes of information (Mayer et al., 2003). Furthermore, this motivation that emerges from learners' social response to the instructional message with social cues yields increases in learning transfer according to the social-agency theory (Mayer, 2014b). As attentional cues, embodied modes such as the instructor's gaze toward the whiteboard guides learners' attention to the relevant part of the multimedia message, which leads to higher cognitive engagement and learning (Stull et al., 2021).

Different experiments that used pedagogical agents with embodied modes in multimedia learning environments provided empirical evidence for supporting social-agency and embodied instruction theories. For example, in Lusk and Atkinson's (2007) study, a fully-embodied pedagogical agent in the form of a parrot gazed in certain directions and utilized gestures to grab learners' attention to the specific parts of learning content about proportion problems. Near and far transfer scores of the students learning with this agent were higher than the scores of the students not seeing the agent. Similarly, Wang et al. (2018) designed an animated pedagogical agent using eye gaze, posture, and gestures to point to the relevant part of the illustration on the board. In two separate experiments, the students learning chemical synaptic transmission with these agents outperformed the control group in retention and transfer tests. Moreover, their eye fixations during the learning process indicated that they allocated more visual attention to the target parts of the illustrations. In a recent study, Schneider et al. (2021) investigated the impact of gestures and facial expressions of a real human instructor in video lectures. Both gestures and facial expressions significantly improved university students' retention and transfer of knowledge about geysers. In addition, the instructor was perceived as the most learning-facilitating and human-like when both embodied modes were available in the video lecture.

Instructional videos

One type of multimedia message is instructional videos that provide words with the instructor's narration and on-screen texts and visuals with slides and animations (Mayer, 2021). As we are in the Internet video age now, instructional videos have become quite accessible in both informal and formal learning settings such as YouTube, Khan Academy, and online college courses (Pi et al., 2021d). Emergency remote teaching during the COVID-19 pandemic has also relied on instructional videos to support learning (Belt & Lowenthal, 2021).

Mayer et al. (2020) categorized instructional videos into four types: (1) lectures that instructors give for online courses, (2) demonstrations in which a procedural task is shown step by step, (3) documentaries in which a narrator describes a film, and (4) shows that display actors' play in an educational episode. To classify instructional videos in terms of their style, Chorianopoulos (2018) proposed two dimensions: human embodiment and instructional media. Human embodiment involves the instructor's hands, talking head, or full body, whereas instructional media includes instruments, whiteboards, slides, animation, or simulations (Chorianopoulos, 2018).

In the design of instructional videos, three demands on learners' cognitive processing of information need to be considered (Mayer, 2021). They are extraneous, essential, and generative processing. *Extraneous processing* is brought by poor instructional design and is not beneficial to supporting the instructional goal (Mayer, 2014a). Emphasizing important materials, placing relevant text and graphics near to each other, and removing extraneous materials are some strategies to reduce the extraneous load (Mayer, 2021). For example, Rodemer et al. (2022) used a dynamic red dot that signaled or highlighted a portion of the visuals corresponding to the content of the narration in an instructional video about chemistry. It was found that students watching the video with this strategy had higher retention scores and visual attention to the content than the control group in the study.

Similarly, Craig et al. (2015) revealed that virtual humans pointing to the relevant items to indicate the narrated content in the instructional video led to higher knowledge retention.

Essential processing emerges in the mental representation of the materials in working memory (Mayer, 2014a). Segmenting videos into small parts, providing pre-training about concepts, and presenting words as narration are suggested to manage essential processing (Mayer, 2021). To illustrate, Biard et al. (2018) used segmentation to divide a video demonstrating a medical and procedural task into sections based on the steps of the task. The group watching the segmented video achieved higher scores in the procedural learning test than the group watching the whole video without any pause. As a result, Biard et al. (2018) inferred that such a segmentation was beneficial for constructing a relevant mental model. Segmentation was also used with the strategies for reducing the extraneous load (e.g., signaling and removing unnecessary content) in Ibrahim et al.'s (2012) study. The combined use of these strategies in the design of instructional videos also resulted in higher retention and transfer scores and lower perceived learning difficulty.

Finally, *generative processing* produced by the motivation to learn enhances the interpretation of the presented materials (Mayer, 2014a). Generative learning activities that require students' active involvement in the learning process can be applied to promote this kind of processing while watching instructional videos. For instance, Fiorella et al. (2020) indicated that the students writing explanations concerning the content of video parts they watched achieved higher retention and transfer scores than those just watching the video lessons about the human kidney. In Pi et al.'s (2021b) study, students watched video lectures teaching some English words in three conditions: (1) watching videos passively, (2) creating sentences with each English word learned from the videos and reading them loudly, and (3) creating sentences and reading them loudly in the presence of a peer. Pi et al. (2021b) found that students had higher learning performance when they created sentences, although they perceived that they invested more mental effort in this condition. The embodiment or showing gesturing instructor is another way to promote generative processing in instructional videos (Mayer, 2021) and is described in the following section.

Embodiment in instructional videos

The embodiment principle in multimedia learning posits that deeper learning occurs "when on-screen agents display humanlike gesturing, movement, eye contact, and facial expressions" (Mayer, 2014b, p. 345). Actual human instructors in videos are also considered pedagogical agents. Similarly, their embodiment in instructional videos involves on-screen instructors' drawing, use of gestures, and maintenance of eye contact (Stull et al., 2021).

Based on empirical evidence, Mayer (2021) proposed four specific embodiment guidelines to enhance learning from instructional videos. First, the dynamic drawing guideline recommends that instructors draw graphics on a board in video lectures rather than using previously prepared graphics (Mayer et al., 2020). Second, gaze guidance favors instructors' gaze shifts between the board and learners rather than continuous and direct gazes at the learners (Mayer et al., 2020). Third, instructors' dynamic gestures are advisable compared to their still appearance (Mayer, 2021). Finally, the perspective principle suggests video lectures with shootings from the first-person perspective rather than the third-person perspective (Mayer et al., 2020). However, these design guidelines were derived mainly from the studies conducted by Richard E. Mayer and his colleagues rather than a comprehensive literature review (Mayer et al., 2020). This scoping review discusses these guidelines by examining all available empirical evidence in the literature in the next sections.

Prior reviews regarding embodiment in instructional videos

There is a lack of a scoping review that investigates embodiment in instructional videos based on an overall analysis of related studies in the literature. Therefore, this study considered prior reviews on the use of nonverbal behaviors of pedagogical agents in multimedia learning environments to gain insights into the potential role of instructors' embodiment in videos. For example, Davis' (2018) meta-analysis study examined pedagogical agents' gesturing and revealed that it significantly contributed to the near transfer of knowledge and retention of learning and decreased cognitive load compared to the static image or no image conditions. However, in a recent study, Castro-Alonso et al. (2021) did not find the moderating effect of the presence of gestures, eye gaze, and facial expression on learning with pedagogical agents in multimedia environments. The authors suggested that agents with and without social nonverbal communication had a similar impact; hence, this result contradicted the embodiment principle. Nonetheless, the findings of these studies on pedagogical agents involving some non-human or humanoid characters with machine-generated voices are inadequate to discuss the design and potential influence of real human instructors' embodiment in videos.

Concerning the use of actual humans, Henderson and Schroeder (2021) reviewed research that investigated the effects of on-screen instructors in videos. They revealed mixed findings regarding their influence on learning and cognitive load, although they found a few studies favoring instructor presence to increase students' satisfaction. In a recent meta-analysis study, Alemdag (2022) found that while instructor-present videos did not significantly affect learning and social presence, they significantly increased students' cognitive load and motivation. However, these studies were only concerned with the presence or absence of instructors in videos. Both Henderson and Schroeder (2021) and Alemdag (2022) suggested a future review study that will examine research on the embodiment principle or instructors' specific behaviors and appearance, elucidate components of embodiments, and synthesize their impact on learning.

Research aim and questions

Instructional videos with on-screen instructors have become prevalent in both formal and informal online education in recent years (Chorianopoulos, 2018; Henderson & Schroeder, 2021). Moreover, there is a burgeoning interest in the research investigating the effect of instructors' embodiment with gestures, eye gaze, and facial expressions in videos (Pi et al., 2021e; Schneider et al., 2021). The lack of a scoping review accumulating and synthesizing relevant studies prevents researchers and practitioners from identifying research trends and gaps (Arksey & O'Malley, 2005; Peters et al., 2015), understanding the design structure of embodied instructional videos, and obtaining design suggestions for an instructional material that evolved into an integral component of diverse online learning settings (Henderson & Schroeder, 2021; Mayer et al., 2020). While multimedia learning principles guide the design of instructional videos, they might not provide fine details about how learning from transitional visual and audio information in a particular type of video can be supported (Fyfield et al., 2022). Hence, there is a call for a review study to unveil available evidence regarding the embodied instructional videos and explain the roles of instructors' gestures and facial expressions or specific forms of instructor presence in learning processes and outcomes (Alemdag, 2022; Henderson & Schroeder, 2021).

Considering the needs in the literature, this study aimed to review empirical research on embodied instructional videos to portray the current state of this field and delve into the design of such videos. It sought answers to the following research questions:

(1) What are the demographics of the research on embodied instructional videos?
(2) What are the video characteristics of the research on embodied instructional videos?
(3) What are the design factors investigated in the research on embodied instructional videos?
(4) What are the outcome variables investigated in the research on embodied instructional videos?
(5) What is known from the research on embodied instructional videos about the effects of their design?

Method

This study is a scoping review that can map key concepts and essential sources and types of evidence of research areas, especially complex and under-explored ones (Mays et al., 2001). Scoping reviews are preliminary efforts before systematic reviews to determine the breadth and depth of a study field, uncover how research on the field is conducted, identify the main characteristics or factors of a concept, summarize available evidence to guide practices, and find knowledge gaps (Arksey & O'Malley, 2005; Munn et al., 2018; Peters et al., 2015).

The scoping review stages proposed by Arksey and O'Malley (2005) and then elaborated by Levac et al. (2010) guided this study. These stages are "(1) identifying the research question, (2) identifying relevant studies, (3) study selection, (4) charting the data, (5) collating, summarizing and reporting the results" (Arksey & O'Malley, 2005, p. 22). The next sections give information about the stages followed after the research questions of this study were determined.

Identifying relevant studies

In the identification of relevant studies for scoping reviews, both comprehensiveness of the review and its feasibility are considered to provide adequate answers to research questions (Levac et al., 2010). For this study, the most used databases containing primary sources in the education field (Frankel et al., 2012) were selected to achieve the research aim. They were Web of Science, Scopus, ERIC, and PsycInfo. The relevant resources were searched in these databases on December 13, 2021. Title, abstract, or keywords fields of the database search engines were filled with the following keywords and Boolean operators: video and (instructor or lecturer) and (face or expression or eye or gaze or gesture or body or movement or drawing or perspective or position). As limiters of the search, peer-reviewed journals were selected to access quality articles. No time span was specified to obtain all available sources. Early-access articles were also included in the results. Finally, the search was limited to English full-text articles.

Study selection

The procedure proposed by preferred reporting items for systematic reviews and meta-analyses guidelines for scoping reviews (Tricco et al., 2018) was followed for study selection (Figure 1). First, articles identified through database search results were listed. This list was enriched with additional studies after the citations used in all selected articles were reviewed, as Peters et al. (2015) suggested. The manual literature search on instructor-present videos in the Google Scholar search engine also provided relevant studies that were incorporated into the list as additional records from other sources. Then, the duplicate articles were removed from the list, which left 777 articles for eligibility assessment. Second, their titles and abstracts were screened to select the relevant ones for this scoping review. Inclusion and exclusion criteria in Table 1 were taken into account to identify them. Consequently, 703 articles were excluded from the review list. Third, full texts of the selected 74 articles were assessed to find the ones meeting the eligibility criteria. This step resulted in 55 eligible articles investigating embodied instructional videos. Figure 1 explains the reasons why the remaining 19 articles were excluded from the review. It is

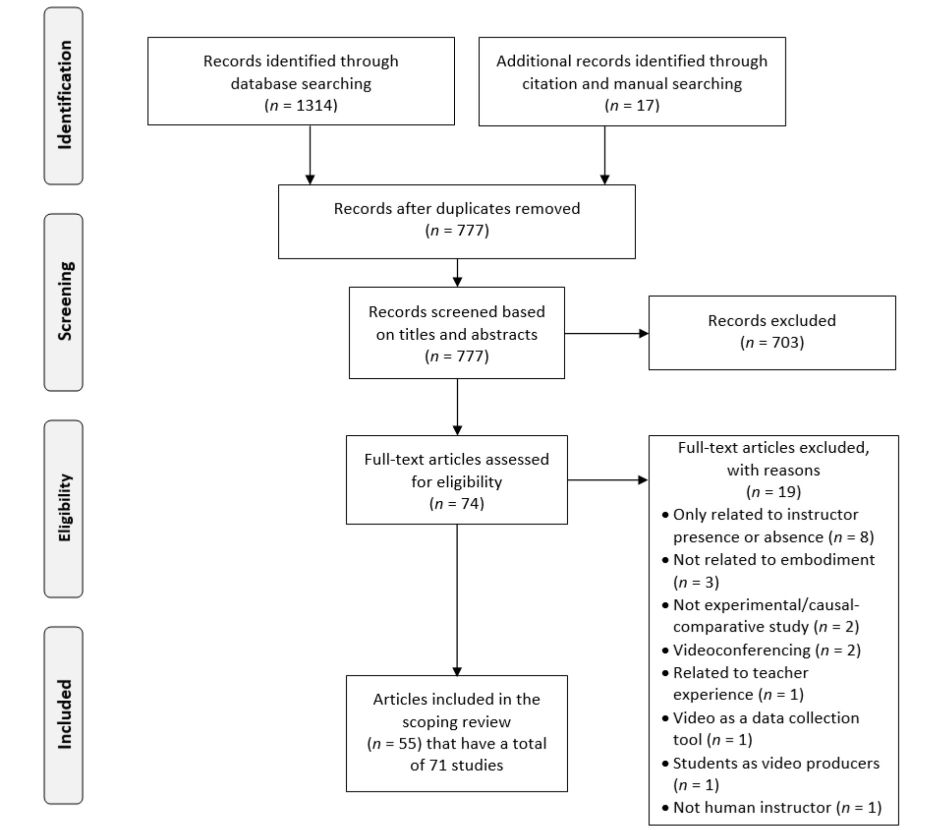

Fig. 1 The flow diagram regarding the study selection

also important to note that 14 articles included more than one experiment or study. Each study in these articles was considered a separate study in the current review; as a result, 71 studies were used for charting and analyzing data and reporting results.

Table 1 Inclusion and exclusion criteria

Criterion	Inclusion	Exclusion
Document type	An article published in a peer-reviewed journal	A non-peer-reviewed article
Type of study	An original and empirical research study	A review, editorial, or conceptual paper
Instructor	A human instructor	An animated character or pedagogical agent
Study focus	Use of embodied modes of information in instructional videos to enhance the learning process or outcomes	Questioning only whether on-screen instructors should be present or absent in instructional videos
Instructional video	Use of instructional videos recorded or produced before the study implementation	Use of student-created videos or videoconferencing technologies for synchronous communication between instructors or learners
Language	Articles in English	Articles not written in English
Availability of full text	Full text available	Full text not available

Charting data

The selected studies were examined to chart or extract data related to (1) their demographics, (2) video characteristics, (3) design factors, (4) outcome variables, and (5) the effects of design factors. A spreadsheet program was used to record them in a table. The additional file (see Supplementary information) provides a copy of the table in which relevant data extracted for each study are presented.

First, the demographics of the studies were examined in terms of the year, learner group, and location of participants. The learner group consisted of K-12 students, higher education students, adults older than 18 using videos for informal education, and older adults about 50 years old.

Second, video characteristics were analyzed in regard to video types, subject domain, instructional media, setting for video watching, average video length, and treatment duration. Prior classifications were used to identify characteristics; however, they were refined based on the analysis results of this study. For example, video types were determined with Mayer et al.’s (2020) categorization of instructional videos. However, a new type (i.e., feedback videos) needed to be added to this categorization, including only instructional videos in lecture, demonstration, documentary, and show types. Feedback videos that provide information about the strengths and weaknesses of a learner’s performance and remedial actions for performance improvement were missing in Mayer et al.’s (2020) categorization. Different from the existing types of videos that present declarative and procedural knowledge to teach a topic, feedback videos give evaluative information about a learner’s performance. Moreover, the former ones are generally designed for all learners in the target group, whereas feedback videos can be more individualized and personalized because they mostly target one specific learner and her/his performance (Henderson & Phillips, 2015). Therefore, it was necessary to extend Mayer et al.’s (2020) categorization with feedback videos. Subject domains of the reviewed videos included science, technology, engineering, math (STEM), and humanities and social sciences, similar to previous reviews (Castro-Alonso et al., 2021; Davis, 2018). This study also added the health category. Analysis regarding instructional media was constructed according to Chorianopoulos’ (2018) classification for video style. It included animation, slides, whiteboards, and instruments. New physical (e.g., model, flipchart, and virtual reality headset) and digital media (e.g., screencast or screen capture) were added to this classification based on the reviewed studies. Finally, time intervals regarding video length were determined following Guo et al.’s (2014) recommendation of using videos less than six minutes.

Third, design factors in embodied instructional videos were determined. Conditions manipulated in experimental studies or groups in causal-comparative studies were checked to find and record a component of design factors. Fourth, to explain the role of design factors in learning from embodied instructional videos, outcome variables in the studies were identified.

Finally, findings regarding the effects of the design factors in the embodied instructional videos were analyzed. To this end, first, the studies were grouped and listed in a new table according to the design factors determined in the findings of the third research question. Then, the main results of the studies available for each outcome variable were written in separate cells in the table. Such a data extraction helped the researcher compare the results of different studies and determine the common or contradictory findings for a specific design factor and outcome variable.

Collating, summarizing, and reporting the results

This stage of the scoping review includes "analyzing the data, reporting results, and applying meaning to the results" (Levac et al., 2010, p. 6). In data analysis, descriptive numerical summary and thematic analysis are applied (Levac et al., 2010). Similarly, this scoping review analyzed study demographics and video characteristics descriptively and summarized them in tables and charts. Moreover, the thematic analysis method was used to collect design factors and outcome variables under overarching themes and provide a better synthesis of findings regarding the effects of embodied instructional videos.

In thematic analysis, the phases proposed by Braun and Clarke (2006) were followed. For example, in the analysis of design factors, each article was read and reviewed several times as a first step for obtaining information about the specific facets of embodiment design in instructional videos. Second, initial codes were generated based on the experimental conditions or groups in the study. Third, the codes were categorized into themes. The data-driven approach was mainly used in the coding. However, in the naming of some design codes and themes, embodiment principles regarding the effective design of instructional videos (Mayer, 2021; Mayer et al., 2020), gesture classifications (McNeill, 1992), dimensions of instructor presence in video courses (Yuan et al., 2021), and design factors of pedagogical agents (Heidig & Clarebout, 2011) were considered to provide themes consistent with the available terms in the literature. Fourth, themes, codes, and related studies were examined together to evaluate their congruity. Finally, each theme was refined and defined. The final set of design factors consisted of seven themes: instructor demeanor, instructor visual presence, learner characteristics, generative activities, content, scene, and instructional media.

As a result of thematic analysis, outcome variables in the reviewed studies were categorized into learning processes measured while learners were studying and learning outcomes measured after studying. The variables regarding learning processes were eye movement, behavior, brain activity, and emotion. Learning outcomes were related to learning, social presence/agency, affect, and perception of the cognitive process.

Findings regarding the effects of embodied instructional videos were also organized thematically according to the seven design factors. All positive, neutral, and negative results were considered with a critical lens. The results supported by multiple studies were reported to reveal the key conclusions about the influence of design factors. Gestures and their effect on learning were examined in a noticeable number of studies (n = 17); hence, gesture types and the number of studies (not) supporting their positive impact were also specified in the analysis.

Finally, intercoder reliability in the data analysis was examined. The author of this research organized a meeting with a doctoral candidate in the instructional technology program to explain the coding list and analyze two studies together. Then, 12 studies (16.90%) out of 71 were evaluated by the doctoral candidate as a second coder. This percentage was higher than 10%, recommended for the sample size in intercoder reliability analysis (Lombard et al., 2002). Cohen's kappa values to measure intercoder reliability were $k_{\text{year}} = 1$, $k_{\text{learner}} = 1$, and $k_{\text{location}} = .89$ for the demographics of the study; $k_{\text{type}} = .75$, $k_{\text{subject domain}} = .54$, $k_{\text{media}} = 1$, $k_{\text{video length}} = .87$, $k_{\text{treatment duration}} = 1$, and $k_{\text{setting}} = .49$ for video characteristics; $k = .83$ for design factors; $k = .76$ for gesture types; and $k = 1$ for the effect of gestures on learning. Except for moderate agreement in the subject domain and setting for video watching, there were substantial and (almost) perfect agreements between the author and second coder in all coding categories according to Landis and Koch's (1977) benchmarks. The disagreements between the coders were also resolved through their discussion. The author took into account the second coder's suggestions and made all necessary revisions by analyzing the studies again.

Findings and discussion

Demographics of the studies

The first research question was related to the demographics of the research on embodied instructional videos. The demographics of the selected studies were analyzed in terms of the year, learner group, and location of participants.

Year

There were 71 studies published between 2006 and 2021. The highest publication occurred in 2021 (n = 21) (Figure 2). Widespread instructional video use to support emergency remote teaching during the COVID-19 pandemic (Belt & Lowenthal, 2021; Mayer, 2021) might have surged research on embodied instructional videos in 2021.

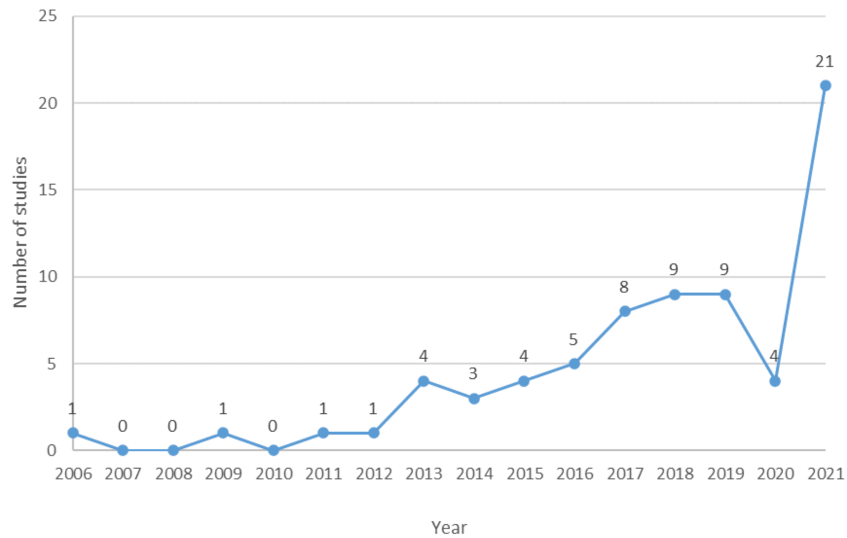

Fig. 2 The distribution of the studies by year

Learner group

Participants of the studies in this review were mainly higher education students (*n* = 65) (Figure 3), similar to the previous reviews on video-based learning environments (e.g., Akçayır & Akçayır, 2018; Birgili et al., 2021). Participation of K-12 students and adults in the studies was noticeably low. Teachers in K-12 education apply flipped and online learning with instructional videos and use videos as in-class materials to support the acquisition of declarative knowledge, show complex phenomena, and increase student engagement and autonomy (Fyfield, 2022). Instructional videos in informal learning

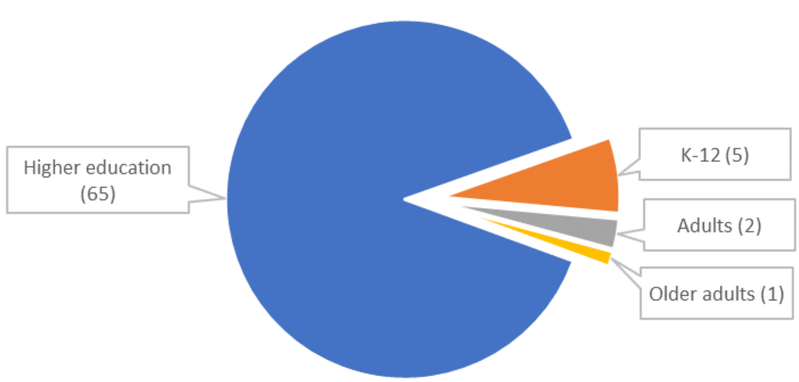

Fig. 3 Learner groups in the studies and their frequencies

Note. Since one study (Ouwehand et al., 2015b) included different learner groups (K-12, higher education, and older adults), the total frequency in the figure exceeds 71.

environments and MOOCs are also a means for adult individuals to pursue life-long learning. Therefore, there is a need for more research conducted with learner groups including K-12 students and adults to investigate how embodiment in instructional videos works for them.

Location of the participants

Students in the reviewed studies were from the USA (n = 34), China (n = 20), Germany (n = 7), Netherlands (n = 5), Spain (n = 2), Canada (n = 1), and France (n = 1). There was one study that did not reveal the location of the participants. The high number of studies of two leading scholars from the USA and China, Richard E. Mayer and Zhongling Pi, brought about such a distribution regarding the locations of the participants. Researchers in other countries need to contribute to the literature in embodied instructional videos to enhance the transferability of findings to learners in different cultures.

Video characteristics of the studies

The second research question was related to the video characteristics of the research on embodied instructional videos. Video characteristics were examined in regard to video types, subject domain, instructional media, setting for video watching, video length, and treatment duration with videos.

Video types

Three video types were identified in the research on embodied instructional videos: lecture (n = 42), demonstration (n = 30), and feedback (n = 2). Some studies (n = 3) included both lecture and demonstration videos. The finding concerning three video types revealed that embodiment in instructional videos could be utilized to provide feedback in addition to its common use for giving a lecture and displaying how to do a procedural task. It also indicated scarce studies on embodied video feedback. Multimodal video feedback including visual, verbal, and gestural information can be more effective in attending to and comprehending feedback content (Henderson & Phillips, 2015; Hung, 2016). However, there is a need for experimental studies that can prove the impact of multimodal video feedback with embodied modes of information.

Subject domains

The subject domains of instructional videos were STEM (n = 43), humanities and social sciences (n = 26), and health (n = 2) (Table 2). The prevalent subjects were related to science topics (n_{biology} = 11, n_{physics} = 8, and $n_{\text{chemistry}}$ =7) and foreign language learning (n = 9). Previous literature reviews on multimedia learning (Alemdag & Cagiltay, 2018;

Table 2 Subject domains of instructional videos and their frequencies

Domain	Subject	*n*
STEM	Biology	11
	Physics	8
	Chemistry	7
	Math	6
	Educational technology	5
	Statistics	3
	Mechanical/material engineering	3
Humanities and social sciences	Foreign language	9
	Geography	5
	Communication	4
	History	3
	Psychology	3
	Sociology	1
	Teaching	1
Health	Nursing	1
	Physical exercise	1

Mutlu-Bayraktar et al., 2019) also revealed more common use of STEM materials, especially science materials, in the reviewed studies. Similar to pedagogical agents, instructors' ability to demonstrate tasks and use gestures to explain or highlight a concept in videos can facilitate the acquisition of abstract science and math constructs and procedures (Schroeder et al., 2013). Likewise, gestures involve a variety of semiotic affordances that support the production, comprehension, and learning of a second language (Gullberg, 2006). Therefore, researchers might have preferred videos on these subjects more frequently.

Instructional media

A variety of instructional media were used in the reviewed studies (Table 3). The most common ones were slides (*n* = 27) and whiteboards (*n* = 16). The animation was used in only one study. The rare utilization of animations in embodied instructional videos raises questions about how students pay attention when instructors use gestures and animations because they are two dynamic and competing visual sources of information for learners. Another relevant finding was the incorporation of screencasts into the embodied instructional videos as media. Two studies in the educational technology domain used screencasts to demonstrate how to use a software program. The instructor appeared near the software image in these videos to explain the procedural information. Screencast videos have become quite prevalent due to most people's preference for watching the demonstrations in these videos rather than reading documents about software usage (Chorianopoulos, 2018). They are also widely used in computer science, especially in

Table 3 Media in instructional videos and their frequencies

Instructional media	*n*
Slides	27
Whiteboard	16
No media	11
Models	7
Instruments	5
Pictures	4
Flipchart	2
Virtual reality headsets	2
Not available	2
Screencast	2
Animation	1

Note. Since some studies (n = 8) used more than one type of media, the total frequency in the table exceeds 71.

programming courses that show the instructor's coding screen (Kokoç, 2019). Similar to the embodied videos with animations, the videos with screencasts can overload students' limited cognitive capacity because of the combined use of dynamic screen activity and moving images of the instructor modeling a task performance (Wouters et al., 2008). Therefore, it is important to conduct more research studies that can guide the instructional design of embodied videos with animations and screencasts for learners not to experience cognitive load while watching these videos.

The setting for video watching

The settings where participants watched instructional videos are provided in Figure 4. In the majority of the studies, settings were laboratory or testing rooms where learners

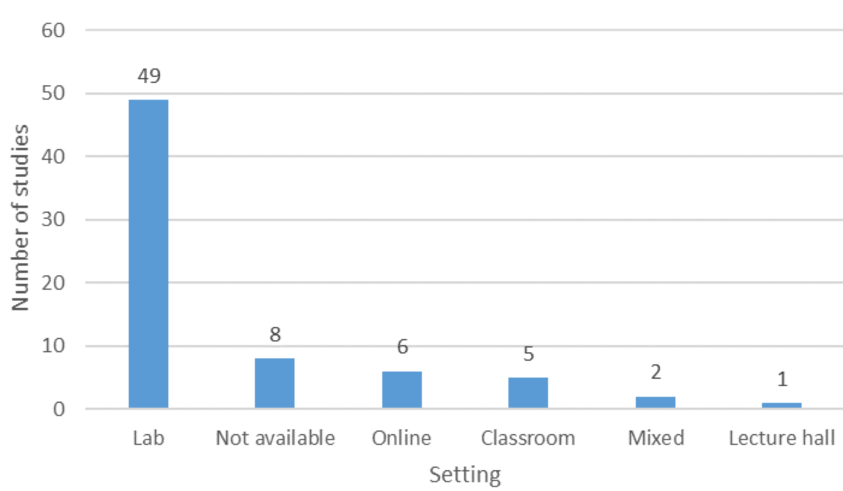

Fig. 4 Settings for watching instructional videos and their frequencies

individually watched the videos (*n* = 49). On the contrary, online environments (*n* = 6), classrooms (*n* = 5), and lecture halls (*n* = 1) were used in fewer studies. Students watch instructional videos online in flipped classrooms, MOOCs, and distance courses at their convenient places. Therefore, the ecological validity of the laboratory experiments on embodied instructional video design might be questionable. As contextual factors can mediate the effect of design principles, there is a need for more studies on instructional videos designed for authentic learning environments rather than laboratory settings (Fyfield et al., 2019).

Video length and treatment duration

Average lengths of instructional videos were generally less than six minutes (*n* = 24) or between six and 12 minutes (*n* = 21) (Figure 5). In regard to treatment duration, instructional videos were presented in one session in most studies (*n* = 65). Other studies allocated two (*n* = 1) and three days (*n* = 1) and one semester (*n* = 4) for the experimental period. Therefore, a limited number of studies investigated the effect of embodied instructional videos in multiple learning sessions or semester-long courses. Presenting instructional videos for a longer experimental period is important to mitigate the novelty effect that arises when the treatment is new and effective only in the initial time (Lodico et al., 2010). It is more likely to occur when new technologies are introduced (Consolvo et al., 2017), such as when virtual reality (VR) is used for embodied instructional videos. However, when the treatment duration lasts several sessions, the learners are accustomed to the new presentation of information, and their attention, efforts, and persistence due to the novelty of the media tend to decrease (Clark, 1983). For example, Merchant et al. (2014) found that learning gains declined when the number of treatment sessions for VR-based instruction with games rose. Therefore, future research can extend the implementation

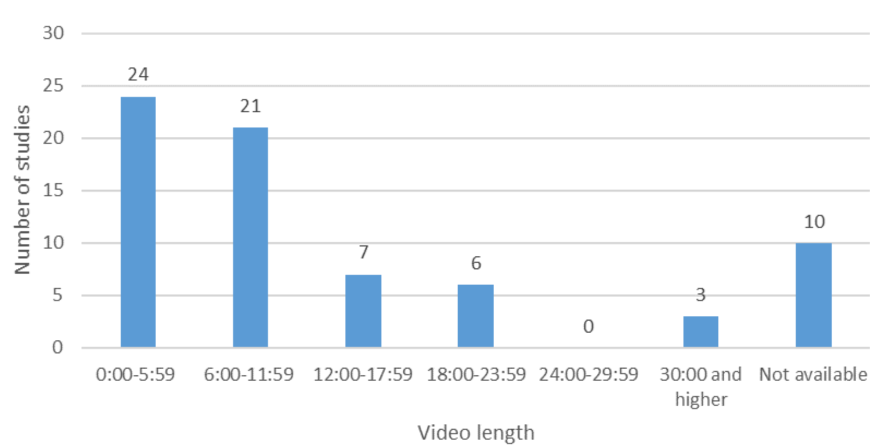

Fig. 5 Average lengths of the instructional videos and their frequencies

period of studies to increase learners' experiences with the treatment and eliminate the initial heightened learner activity due to the novelty effect.

Design factors in the studies

The third research question was related to the design factors investigated in the research on embodied instructional videos. Design factors in the selected studies were categorized into seven themes in this review: instructor demeanor (*n* = 43), instructor visual presence (*n* = 15), generative activities (*n* = 10), learner characteristics (*n* = 10), content (*n* = 8), scene (*n* = 8), and instructional media (*n* = 8). Figure 6 displays codes under each theme with their frequencies. This section describes each theme, but the findings regarding the fifth research question give more information about these factors while explaining their effects on learning processes and outcomes.

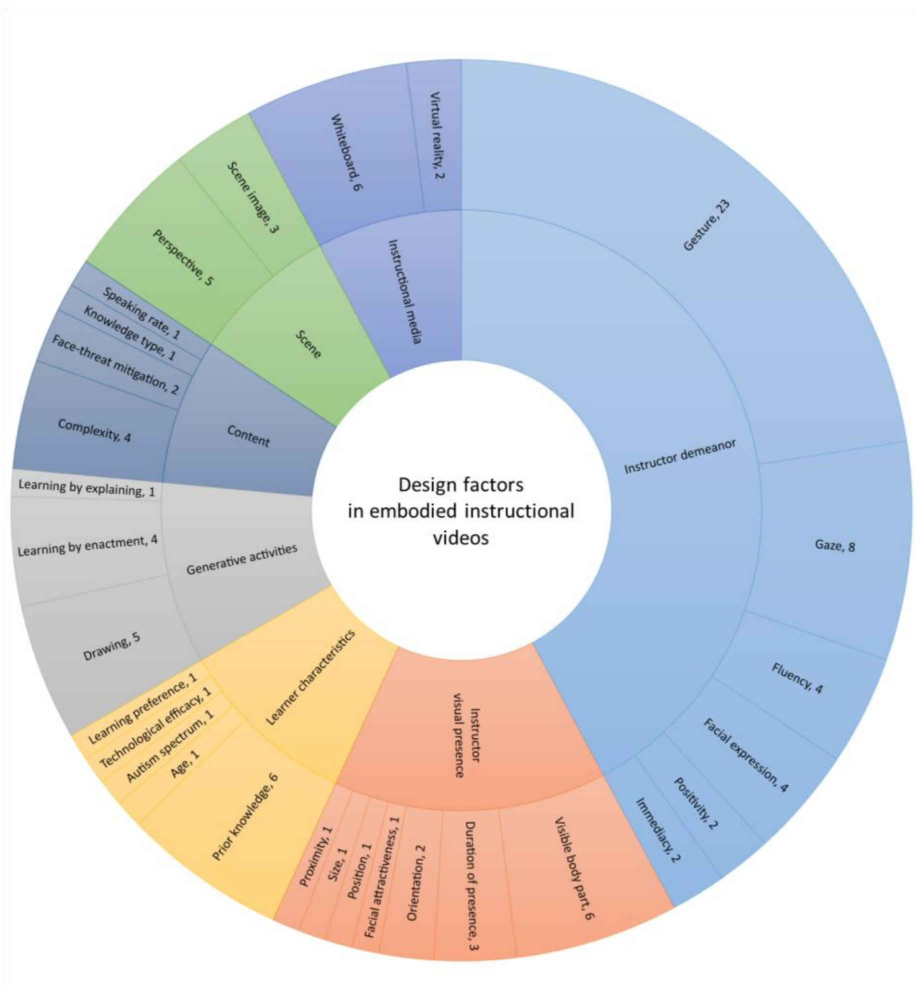

Fig. 6 Design factors in the studies on embodied instructional videos

The first design factor is instructor demeanor. It is related to how instructors behave and look in instructional videos. Instructors' gestures, gaze, facial expressions, and fluency of their behaviors are some specific embodied cues regarding instructor demeanor (e.g., Carpenter et al., 2013; Ouwehand et al., 2015a, Pi et al., 2017a, 2019a, 2019b; Schneider et al., 2021; Stull et al., 2021; Wang et al., 2019b). Second, instructor visual presence refers to how the instructor's appearance is formatted or organized in the video scene, such as in terms of orientation, position, size, and duration (e.g., Pi et al., 2017b; van Gog et al., 2014; Yi et al., 2019; Yu et al., 2021). Third, learner characteristics involve consideration of learners' demographics and cognitive and motivational characteristics in the design of embodied instructional videos (e.g., Fiorella & Mayer, 2016; Lyons et al., 2012; Pi et al., 2019a). Fourth, generative activities make instructors or learners physically and cognitively productive during the video (e.g., Fiorella & Mayer, 2016; Fiorella et al., 2017, 2019). Fifth, the content of embodied instructional videos is related to the nature of the content (e.g., complexity and knowledge type) and how texts and visuals are organized (e.g., Fiorella et al., 2017; Gluhareva & Prieto, 2017; Pi et al., 2021a). Sixth, the scene is a design factor that deals with how the video environment appears from the learners' perspective (e.g., Boucheix et al., 2018; Jackson et al., 2006; Merkt et al., 2019). Finally, instructional media such as whiteboards and VR headsets are the tools instructors or learners use during video recording or watching (e.g., Hekele et al., 2021; Stull et al., 2018b, 2021).

Prior review studies questioned whether instructor presence is beneficial for learning from videos (Alemdag, 2022; Henderson & Schroeder, 2021) and focused on only instructors' dynamic drawing, gaze guidance, gestures, and camera perspective (Mayer, 2021). There is a lack of study that indicates components of embodiment in instructional videos (Henderson & Schroeder, 2021). This study provided a rich list regarding the design factors of embodied instructional videos based on a scoping review of the literature. It indicated that the development of embodied instructional videos is a more complex design and research area. Instructor movements, appearance, and use of instructional media need to be considered from different aspects with content elements and learner characteristics, activities, and tools.

These results imply the importance of taking a broader perspective in the design of embodied videos. They indicate that the design elements of these videos also involve instructors' facial expressions, fluency, positivity, amount and duration of visibility, and use of whiteboards. Moreover, the current study emphasizes analyzing learners, content, and media while selecting the embodiment as an instructional strategy. Such an analysis can give information about which type of embodied cues might be more effective in a particular learning context. Overall, this scoping review extends Mayer's (2021) embodiment principle by uncovering all design elements available in the literature.

Researchers can also benefit from these design factors to provide detailed descriptions of their embodied instructional videos and conditions under which participants watch them in their studies. As a result, findings of similar research on instructional videos can be compared considering the learning materials and conditions, and meaningful replication and meta-analysis studies can be conducted (Fyfield et al., 2022). However, in a recent systematic review, Fyfield et al. (2022) found that most studies had limited information about the instructional videos used in the research. This review guides future studies by presenting a comprehensive list that can indicate what to include in the description of embodied instructional videos.

Outcome variables in the studies

The fourth research question was related to the outcome variables investigated in the research on embodied instructional videos. Outcome variables in the selected studies were collected under two main themes in this review: learning processes (*n* = 23) and outcomes (*n* = 164) (Figure 7).

Learning processes

Measures regarding learning processes were related to learners' eye movement (*n* = 16), behavior (*n* = 3), brain activity (*n* = 3), and emotion (*n* = 1) while they were studying instructional videos. First, eye movement measures included first fixation time to the relevant content, number and duration of fixations on instructor and content area, and

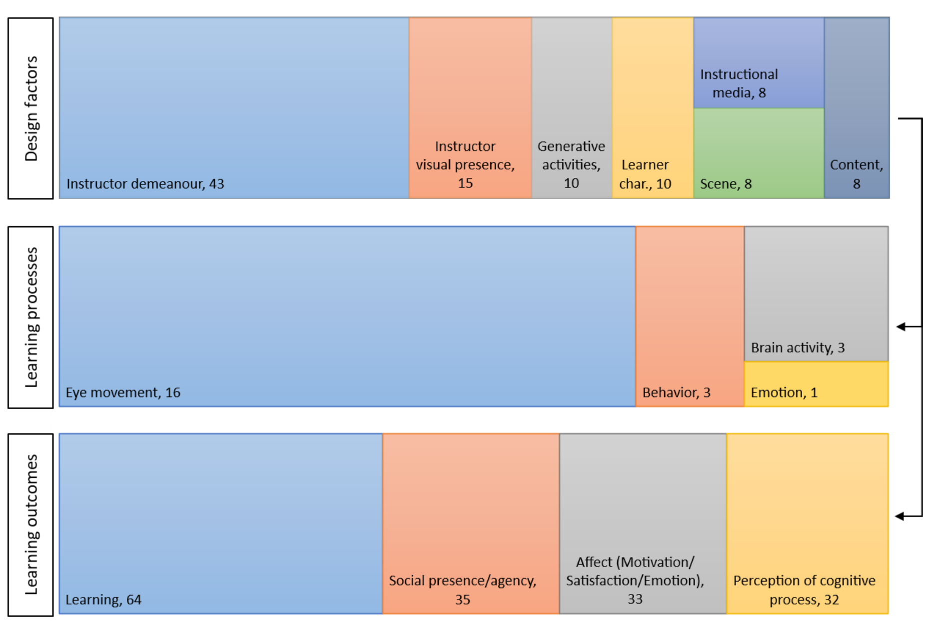

Fig. 7 Design factors with related learning processes and outcomes

transitions between instructor and content. They were used to evaluate learners' visual search efficiency, cognitive processing of multimedia elements, and visual attention distribution in instructional videos (e.g., Pi et al., 2021c, 2021d; Stull et al., 2021). Second, behavioral measures consisted of time for restudying video content before testing (Carpenter et al., 2013, Experiment 2), the correctness of imitation behaviors during video watching (Jackson et al., 2006), and attrition and assessment taking rates in video lectures (Kizilcec et al., 2015, Experiment 2). Third, learners' brain activities were monitored through functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) devices to map the activation of brain regions and measure the amplitude of alpha and beta powers during learning with different video designs (Jackson et al., 2006; Pi et al., 2021a, Experiment 1; Pi et al., 2021e, Experiment 2). Fourth, one study (Wang et al., 2019a) tracked learners' facial expressions during video watching to determine the arousal level of their emotions.

Learning outcomes

In terms of learning outcomes measured after studying, there were four main categories: learning (n = 64), social presence/agency (n = 35), affect (motivation/satisfaction/emotion) (n = 33), and perception of cognitive process (n = 32). First, learning outcomes were assessed with objective test items regarding retention, transfer, comprehension, and problem-solving (e.g., Fiorella & Mayer, 2016; Koumoutsakis et al., 2016; Pilegard & Fiorella, 2021), observation of demonstration performance (e.g., Fiorella et al., 2017, Experiments 1 & 2), and questionnaires regarding perceived learning (e.g., Carpenter et al., 2013, Experiments 1 & 2). Second, social presence/agency referred to learners' feeling of presence in the learning environment created by instructional videos and their perceptions of instructors' characteristics and interactions with learners (e.g., Beege et al., 2020, Experiments 1 & 2; Wang et al., 2019b; Witt & Kerssen-Griep, 2011). Third, affective outcomes involved measures of motivation, satisfaction, and emotions felt for learning with specific designs of instructional videos (e.g., Carpenter et al., 2013, Experiments 1 & 2; Waller et al., 2021; Yuan et al., 2021). Fourth, measures regarding perceived mental effort, mental load, difficulty, and attention during learning with embodied videos provided inferences about students' perceptions of cognitive process (Beege et al., 2020, Experiments 1 & 2; Ouwehand et al., 2015a, 2015b; Yi et al., 2019).

Compared to the learning outcomes, learning processes have been less investigated in the reviewed studies. Learning process measures include neurocognitive and psychophysiological measures (e.g., eye movements and brain activity) that can overcome the subjectivity of self-rating measures and provide inferences about cognitive functioning based on physiological states during video watching (Kruger & Doherty, 2016). Such measures are recommended to examine the attentional and cognitive processes and identify the specific instructor behaviors that can enhance learner engagement (Wang et al., 2020). Considering the increasing utilization of these measures in the multimedia learning field (e.g., Alemdag & Cagiltay, 2018; Ozel et al., 2021), future studies can prefer analyzing the influence of design factors in embodied instructional videos with both learning outcome and process variables to provide multiple types of evidence for design suggestions.

The influence of design factors in embodied instructional videos

The fifth research question was related to the findings regarding the effects of the design of embodied instructional videos. This section explains each design factor displayed in Figure 6 and its potential impact on the learning process and outcomes by providing examples from the reviewed studies. As a result, design considerations are proposed based on the prevalent study findings.

Instructor demeanor

Only instructor presence in videos might not be adequate to enhance learning (Mayer et al., 2020). Indeed, social cues regarding instructor demeanor such as gestures, gaze, and facial expressions are necessary to direct students' attention to relevant content and promote learning (Polat et al., 2021). In this scoping review, the most investigated factor regarding instructor demeanor was gestures. There were 17 experimental studies that compared the effects of gesture-absent and gesture-present instructional videos on learning. The 13 studies (Table 4) reported at least one positive and significant effect of a specific type of gesture in a learning test or condition.

Table 4 The frequencies regarding the effect of specific gestures on learning and related studies

	Gesture				
Effect	Deictic	Beat	Metaphoric	Iconic	Mixed
Positive effect	7 (Beege et al., 2020, Experiments 1 & 2; Koumoutsakis et al., 2016; Pi et al., 2017a, 2019a, 2019b, 2021a, Experiment 2)	3 (Gluhareva & Prieto, 2017; Pi et al., 2021d, Experiment 2; Pi et al., 2021e, Experiment 1)	1 (Schneider et al., 2021)	1 (Carlson et al., 2014)	1 (Pi et al., 2019a)
No positive effect	4 (Fiorella & Mayer, 2016, Experiment 1; Ouwehand et al., 2015a, 2015b; Yeo et al., 2017)	3 (Beege et al., 2020, Experiments 1 & 2; Pi et al., 2021d, Experiment 1)	-	-	-

Note. Three studies investigated the effects of more than one type of gesture.

Instructors' dynamic gestures are recommended by Mayer (2021), but what type of gestures promote learning from videos is unknown. Therefore, gesture types found beneficial for learning in these studies were also analyzed to determine the conditions of this recommendation. According to McNeill's (1992) classification, they were deictic gestures that indicate or point to an object, event, or direction (n = 7), beat gestures or rhythmic movements of hands (n = 3), metaphoric gestures that convey an abstract metaphor (n = 1), iconic gestures that represent a concrete object or action (n = 1), and the mixed group including both metaphoric and iconic gestures (n = 1) (Table 4). Despite some conflicting results, deictic gestures mostly contributed to the learning in the relevant studies. In addition, the positive effect of deictic gestures was noted in regard to students' eye movements during the learning process. Four studies (Ouwehand et al., 2015a; Pi et al., 2017a, 2019a, 2019b) revealed that learners used less time to find the visual area the instructor addressed and dwelled longer on content in instructional videos with deictic gestures. The positive impact of such gestures on attention allocation and learning outcomes was also pronounced in the research on multimedia learning environments with animated pedagogical agents, especially when deictic gestures pointed to the specific visual area corresponding to the narration (Li et al., 2019). This type of gesture provides signaling or cueing that highlights where to look at visual content (Li et al., 2019). Overall, these results advocate the use of dynamic gestures to enhance learning similar to Mayer's (2021) embodiment principle, but they also draw attention to the gesture type. Based on the available evidence in the literature, this review can highlight pointing gestures to enhance learning performance, visual search efficiency, and attention to content during the learning process. More research studies are needed to investigate the effects of other types of gestures and their boundary conditions.

In terms of the influence of instructors' *gaze*, some findings revealed that fixed or direct gaze could cause less attention to the content area but more attention to the instructor than guided, shifted, or directed gaze during the learning process with instructional videos (e.g., Pi et al., 2019b, 2020a, 2020b; Wang et al., 2019b). For example, Wang et al. (2019b) investigated the effect of gaze guidance when the instructor switched her looking between the camera and content area in slides in declarative and procedural knowledge types. They found that the learners allocated more visual attention to the content in both knowledge types, made fewer transitions between the content and instructor, and attained higher social presence and learning in instructional videos with gaze guidance. There were also contradictory results in the reviewed studies. Stull et al. (2021) revealed that the instructor's fixed gaze at the only camera was more favorable for learners' attention toward the content on whiteboards than gaze shifts, and there was an interaction effect between whiteboard type and gaze behavior in terms of learners' attention to the instructor. Moreover, Pi et al. (2021c) indicated that learners watching videos with the instructor's direct gaze and surprised face allocated more attention to instructional slides than those watching videos with gaze guidance. On the whole, this study can provide some evidence supporting Mayer et al.'s (2020) gaze guidance principle. However, considering the conflicting findings, more research studies are needed to determine when gaze guidance is more beneficial for both attentional distribution and learning. The effect of the instructors' gaze might change based on instructional media and other embodied cues in the video.

Another embodiment cue related to instructors' demeanor was *facial expressions* that were not indicated in Mayer's (2021) embodiment principle for instructional videos. Specific facial expressions are easily identifiable cues of certain emotions (Carroll & Russell, 1996). Emotional designs in multimedia learning were found effective in enhancing learning outcomes, positive affect, intrinsic motivation, and mental effort in the recent meta-analysis by Wong and Adesope (2021). Except for Pi et al.'s (2021c) study that favored the neutral face rather than the surprised face in the instructional videos with guided gaze, three studies in this scoping review also revealed the positive impact of facial expressions on learning. Happy face promoted learning in the direct gaze condition (Pi et al., 2020a); facial expressions enhanced retention and transfer in the videos both with and without gestures (Schneider et al., 2021); and the group with a heightened level of expressiveness outperformed conventional level group on medium-level recall (Wang et al., 2019a).

Instructors' demeanors were also investigated in terms of *fluency, immediacy, and positivity*. The common finding was that they led to a more positive evaluation of instructors as social agents. Fluent instructors were perceived as more organized, knowledgeable, prepared, and effective (Carpenter et al., 2013, Experiments 1 & 2; Toftness et al., 2018, Experiments 1 & 2). Instructors with high immediacy were deemed more competent and caring (Kerssen-Griep & Witt, 2015; Witt & Kerssen-Griep, 2011). Moreover, the instructors with positive emotions were appraised as more credible, engaging, and facilitating (Lawson et al., 2021, Experiments 1 & 2). In addition to instructors' gestures, facial expressions, dynamic drawings, and use of objects and contents in videos, natural and frank instructor behaviors are prioritized by learners to enhance the effectiveness of instructional videos (Fidan & Debbag, 2022). The findings of this scoping review based on students' perceptions also imply that fluent and intimate instructors can be more favorable social agents in instructional videos.

Instructor visual presence

Instructor visual presence referred to the arrangement of on-screen instructors in terms of the visible body part, duration of presence, orientation, facial attractiveness, position, size, and proximity. Concerning visible body parts, five studies investigated whether it is necessary to display the instructor's face or body in addition to hands in demonstration videos. Fiorella and Mayer (2018) assert that mere instructor presence might interfere with learners' focus on the content of the instructional message and cause extraneous cognitive load. In contrast, it can also promote a social connection between instructors and learners (Fiorella & Mayer, 2018).

In line with the interference hypothesis, van Wermeskerken and van Gog (2017) and van Wermeskerken et al. (2018) found that learners paid less attention to the demonstration area in the videos when the instructor's face was visible. Moreover, face visibility neither enhanced nor hindered learning in the first viewing of instructional video in other studies (van Gog et al., 2014; van Wermeskerken & van Gog, 2017; van Wermeskerken et al., 2018). On the contrary, using only the hand as an embodiment cue was found more effective in enhancing the transfer of learning than the appearance of the full body (Fiorella & Mayer, 2016, Experiment 4). Therefore, showing only the instructor's hand might be adequate to use embodiment in demonstration videos if other body parts do not have a specific visual function.

The optimal duration of instructor presence was also investigated in three studies. The two studies (Yi et al., 2019; Yu et al., 2021) favored intermittent instructor presence to improve learning gains compared to continuous teacher presence. However, there were contradictory findings concerning cognitive load. While Yi et al. (2019) revealed lower cognitive load in intermittent presence, Kizilcec et al. (2015) and Yu et al. (2021) indicated higher cognitive load when instructors appeared at certain times. Therefore, it is essential to note that while intermittent instructor presence can be more beneficial for learning than continuous presence, it might cause more cognitive load. However, there is a need for more studies to determine when and how long the instructors should appear in videos to enhance learning without imposing cognitive load.

Learner characteristics

It is essential to determine the boundary conditions for whom instructional videos are more beneficial for learning (Fiorella & Mayer, 2018; Mayer, 2021). Learners' cognitive, metacognitive, motivational, and emotional characteristics can mediate the impact of pedagogical agents on learning (Heidig & Clarebout, 2011). In this scoping review, the significant main effect of prior knowledge on learning retention or transfer was found in four studies (Fiorella & Mayer, 2016, Experiments 1, 2, 3, & 4). The remarkable finding in two studies (Fiorella & Mayer, 2016, Experiment 1; Pi et al., 2019a) was that the learners with a low or medium level of prior knowledge benefited from instructional videos with dynamic drawing and gestures more to enhance learning transfer, although there was no significant difference for the learners with high prior knowledge. These findings support the expertise reversal principle in multimedia learning, positing that design guidelines effective for novice learners might not be effective for learners with high prior knowledge

(Kalyuga, 2014). Concerning other learner characteristics, it was indicated that less technological efficacy was related to less perceived learning from instructor-present videos (Lyons et al., 2012). Older adults were also disadvantageous in learning from instructional videos compared to young adults (Ouwehand et al., 2015b). Therefore, learner characteristics, especially prior knowledge, can be a critical moderator of the effect of embodied instructional videos on learning.

Generative activities

Generative learning refers to learners' active involvement in making sense of instructional material by selecting, organizing, and integrating incoming information (Fiorella & Mayer, 2015). One of the generative learning strategies is learning by the enactment related to students' task-related movements during learning (Fiorella & Mayer, 2015). Concerning this strategy, there were findings supporting learners' imitation of models during or after watching demonstration videos to enhance learning (Fiorella et al., 2017, Experiment 1; Stull et al., 2018c, Experiments 1 & 2). In addition, Jackson et al.'s (2006) MRI study revealed that learners' imitation of physical movements in the instructional video led to higher signal change in the extrastriate body area mainly activated in the visual perception of body parts.

This review also considered some instructor movements (e.g., drawing) as generative activities because instructors become physically active to produce learning content while the video is recorded. Three studies (Fiorella & Mayer, 2016, Experiments 1 & 2; Fiorella et al., 2019, Experiment 1) reported that instructors' drawing on a whiteboard while explaining or demonstrating a science topic led to higher retention or transfer scores than showing already-drawn visuals. Moreover, in one study (Swenson et al., 2021), most students favored hand-written content rather than pre-written content. Signaling, temporal contiguity, and segmenting principles of multimedia learning explain why observing instructor drawing can be better for learning (Fiorella & Mayer, 2016). Instructors' hand guides learners' attention to the relevant visuals (signaling) that are drawn synchronously with verbal information (temporal contiguity) and shown one by one at a time (segmenting). Overall, it might be inferred that both instructors' and learners' generative activities in demonstration videos can promote learning.

Content

The content of instructional videos also varied in nine studies in terms of task or visual complexity, face-threat mitigation, knowledge type, and speaking rate. The complexity of the multimedia material moderates the effort allocated for processing and integrating textual and visual information (Schmidt-Weigand et al., 2010). Correspondingly, the moderating effect of content complexity on learning from embodied instructional videos was noted in four studies about gestures and the first-person perspective. Pi et al. (2021d, Experiment 2) found a positive impact of instructors' head nods and beat gestures on retention and transfer performance in instructional videos with simple visual complexity. On the other hand, three studies determined that first-person perspective (Fiorella et al., 2017, Experiment 1), beat gestures (Gluhareva & Prieto, 2017), and pointing gestures (Pi et al., 2021a, Experiment 2) had a significant effect in instructional videos with tasks in high visual complexity or task difficulty. Instructional design becomes more critical when complex materials impose a high working memory load because of their intrinsic nature (Pass & Sweller, 2014). The use of multimedia design principles in the design of such materials can help learners not to exceed their working memory capacity to process and learn the information, although it might not be urgent to utilize the principles in materials with light intrinsic cognitive load (Pass & Sweller, 2014). Therefore, it appears that the design of instructional videos with beneficial embodiment can be more effective for learning content with high complexity.

Scene

The use of cameras is advantageous for displaying a task from different viewpoints, which might not be possible in face-to-face settings (Fiorella & Mayer, 2018). Viewpoints of videos or camera perspectives were first-person, third-person, or mixed in the reviewed studies. Viewing instructors' behaviors from the first perspective in demonstration videos was found more effective in enhancing learners' performance and completion times in circuit building and physical movement tasks than viewing from the third-person perspective (Fiorella et al., 2017, Experiments 1 & 2; Jackson et al., 2006). Jackson et al.'s (2006) MRI study also indicated a more direct mapping of behaviors in the first perspective without further visuospatial transformation. Unlike previous studies, Boucheix et al. (2018) investigated the impact of mixed-use of two perspectives in the video that demonstrated how to insert an indwelling catheter in a patient to better mimic real-life learning situations in nursing education. They found that learners' demonstration performance in mixed perspectives was higher in the steps that required multiple views than the performance of the first-person perspective and control groups. Based on these findings, this review corroborates Mayer's (2021) perspective principle and proposes that demonstration videos should include shootings from the first-person perspective. It saves learners from transforming the representations in the third perspective into their own perspective (Fiorella et al., 2017). First-person perspective in instructional videos can also help learners feel as they perform actions shown in the video and become more immersed in the learning environment, as in digital games and virtual reality applications (e.g., Denisova & Cairns, 2015; Gorisse et al., 2017). However, it might be better to alternate between the first-person and third-person perspectives to show actions from different viewpoints in instructional videos.

Instructional media

Two new technological tools were used in the reviewed studies as instructional media. The first one is transparent videos that provide access to the instructor's face while writing and drawing (Stull et al., 2021). The educational affordances of these videos enable learners to see the full content on the board easily with social cues and gaze guidance (Stull et al., 2018a). The second one is VR which shows 360° videos for viewers to see everything in the range of camera perspective via special headsets and feel more immersed and present in the learning environment (Violante et al., 2019). Six instructional media studies compared videos with conventional or transparent whiteboards, and two of them compared two-dimensional and VR videos. The main effect of whiteboards on learning was reported in three studies favoring the transparent whiteboard (Fiorella et al., 2019, Experiment 2; Stull et al., 2018a, Experiments 1 & 2). However, three studies (Hekele et al., 2021; Stull et al., 2018b, 2021) failed to indicate the main effect of the whiteboard or VR. Furthermore, the results regarding learners' motivation and feeling of social partnership with instructors were inconsistent in the reviewed media studies, and the new media did not lead to a significant decrease in mental load. Therefore, the effect of instructional media to improve embodiment and enhance learning from instructional videos is equivocal. According to Clark (2001), who advocates the significant impact of the instructional method on learning, future studies can investigate the role of media in enhancing cognitive efficiency such as learning time.

Overall, this study analyzed how design factors concerning the embodied instructional videos influenced learning processes and outcomes by reviewing empirical research in the literature. Figure 8 provides a summary of the findings. In particular, it suggests specific design considerations that can guide the development and use of embodied instructional videos to enhance learning.

Mayer (2021) recommended showing gesturing instructors through the embodiment principle based on the results of his empirical studies. A recent systematic review by Fyfield et al. (2022) also corroborated the use of the embodiment principle to enhance learning from instructional videos as a result of the analysis of 11 studies. However, there was a lack of a comprehensive review study concentrating on embodiment in instructional videos and its effect on learning (Henderson & Schroeder, 2021; Mayer et al., 2020). Moreover, boundary conditions under which embodiment instructional videos work in an optimal way are unclear (Doherty, 2022; Fiorella & Mayer, 2018; Mayer, 2021). This scoping review examining 71 studies provides detailed information about the influence of specific embodied cues and learning conditions, as shown in Figure 8. Although there is a

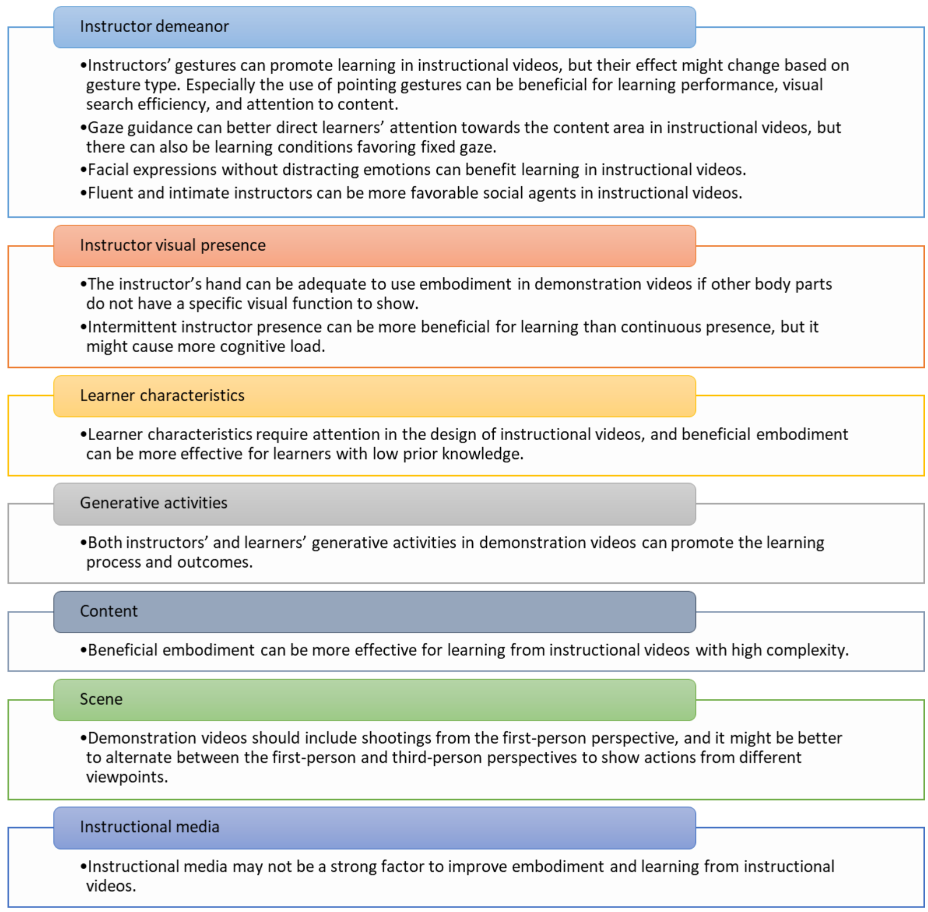

Fig. 8 Summary of the findings regarding the influence of design factors in embodied instructional videos

need for more empirical studies testing the validity of these design considerations, they can be helpful for practitioners and researchers to develop effective embodied instructional videos.

Practical implications

This study reviewed research on embodied instructional videos to analyze the characteristics of such videos, determine their design factors, and synthesize the prevailing results regarding the influence of their design on the learning process and outcomes. As a result, seven design factors were identified: instructor demeanor, instructor visual presence, generative activities, learner characteristics, content, instructional media, and scene. These diverse factors imply a need for a comprehensive analysis of design cases for embodied instructional videos. In particular, practitioners need to consider both instructors' embodiment and learners, content, and context while deciding how to enhance the effect of instructional videos on the learning processes and outcomes with the embodiment. A detailed list of design considerations related to seven factors and their influence on learning processes and outcomes is provided in Figure 8.

The key findings regarding the impact of embodied video design support four embodiment principles: dynamic drawing, gaze guidance, dynamic gestures, and the first-person perspective (Mayer, 2021; Mayer et al., 2020). This review also brought new considerations to enhance learning with specific design decisions. These considerations suggest instructors' use of deictic gestures, facial expressions without distracting emotions, and fluent and intimate behaviors. In addition, the display of only instructors' hands in demonstration videos can be recommendable if other body parts do not have a specific visual function. Furthermore, students' imitation of the models in demonstration videos is advisable to promote generative learning. Finally, learners with a low level of prior knowledge and subjects with high task and visual complexity can be important conditions in which embodied instruction in videos can have a greater influence on learning. Previous experiments support the higher effect of some embodied cues such as gestures on low knowledgeable learners (e.g., Fiorella & Mayer, 2016, Experiment 1; Pi et al., 2019a) and complex materials (Gluhareva & Prieto, 2017; Pi et al., 2021a, Experiment 2).

Recommendations for future research

This study also has several recommendations for future research. First, as most of the reviewed studies were conducted with higher education students, researchers can consider other learner groups in the subsequent studies. For example, they can involve K-12 education students or older adults to discuss the generalizability of the existing findings and determine the specific needs and capabilities of a learner group regarding embodied instructional videos.

Second, concerning video characteristics, video lectures and demonstrations were common instructional video types. However, video feedback was identified in a few studies. Future research can use video feedback to indicate how learners build information on the product they generated with instructor embodiment. Moreover, the least selected instructional media in the reviewed studies were animations and screencasts. It is suggestible for the researchers to investigate the effect of these media because they generally provide transient pieces of information which might be more difficult for learners to follow while watching embodied modes of communication. In addition, it is recommendable to use embodied instructional videos in real learning settings for a more extended period to enhance the ecological validity of the findings and prevent novelty effects.

Third, regarding design factors, instructor demeanor and visual presence were the most investigated factor. The presence/absence of specific cues regarding these factors was the common research concern in the reviewed studies; on the contrary, the combination of the instructors' embodied cues has been analyzed in fewer studies. As real teaching requires the use of different embodied cues (Pi et al., 2021e), it is necessary to integrate more than one cue in future instructional videos.

Fourth, it is also possible to see symbolic cues such as colored boxes and words to signal or highlight a portion of visual content in addition to embodied cues in instructional videos. For example, Moon and Ryu (2021) utilized both a pedagogical agent's gestures and colored boxes in video instruction. In such a case, they found lower learner attention on the content area and lower learning comprehension. Therefore, prospective studies need to investigate the combined effect of instructors' embodied cues and symbolic cues on the learning process and outcomes to elucidate the interplay between the cues in instructional videos.

Fifth, there were design factors pertaining to learners, content, and context; however, the number of studies that examined the interaction effect of these factors and instructor demeanor and presence was limited. This situation prevents the researchers from providing design recommendations with boundary conditions. Therefore, it is critical for future studies to consider for whom, what, and in which context a particular instructional video design can benefit learning. The use of learning process measures is also suggestable to provide a more fine-grained analysis of learning events in embodied instructional videos.

Sixth, this study reviewed research on instructional videos that provided asynchronous lectures, demonstrations, and feedback. Another common way to deliver video lectures in online and blended education is by using video conferencing technologies (Belt & Lowenthal, 2021). These technologies enable synchronous interaction among instructors, students, and content, and students can demonstrate their use of embodiment with their cameras. Because of the aforementioned affordances of video conferencing technologies, the role and influence of instructors' embodiment might be different in synchronous video lectures. Future studies can investigate embodied instruction in these learning environments in detail.

Finally, embodied instructional videos can also be used for fostering online discussions, although this scoping review did not find a related study. For example, instructors' gestures can highlight what students should pay attention to while discussing a topic. Facial expressions might indicate instructors' emotions and reactions concerning the quantity and quality of online discussions. The use of instructional videos can promote students' interaction, social and teaching presence, and positive perceptions regarding online discussions (Belt & Lowenthal, 2021). However, there is a need for studies that can provide empirical evidence for the potential benefits of embodied instructional videos in online discussions.

Limitations

There are some limitations of this research. First, only English articles published in peer-reviewed journals were included in this review. Conference papers and theses in other languages can also be examined to increase the research scope in future studies.

Second, this scoping review indicated that some authors contributed to the research on embodied instructional videos with several studies. While this situation increased the number of related studies, it might have caused biased results due to the increased possibility of the participation of learners with similar characteristics, the use of the same learning materials, and the measurement of the same design, process, and outcome variables. The current study invites different researchers to conduct replication studies in a variety of contexts and investigate the less explored study areas of embodied instructional videos to widen the breadth of this field.

Third, an overall effect size was not calculated in synthesizing the results regarding the effect of instructors' gestures because scoping review studies aim to give a general overview of the available evidence in a broad research field rather than investigating the effectiveness of a particular treatment or practice (Peters et al., 2015). However, scoping reviews can indicate whether there is adequate evidence in the literature that can answer the questions of the prospective systematic reviews (Munn et al., 2018). Considering the number of sources identified in this review, the following studies can perform a meta-analysis and analyze how the effects of different types of gestures change as a function of study-level covariates.

Finally, the design considerations proposed in this study were based on the prevalent findings regarding the influence of design factors in the reviewed studies. More research is needed to obtain more robust conclusions by systematically analyzing the effect of a specific design factor on an outcome variable. This scoping review exploring the nature of evidence in the field of embodied instructional videos can be a precursor to future systematic reviews (Munn et al., 2018).

Conclusion

In parallel with the regaining popularism of instructional videos in formal and informal learning environments in the last years, the number of research studies increased to determine when an instructional video results in better learning (de Koning et al., 2018). One proliferating research area in instructional video design has been the use of embodiment (Pi et al., 2021e; Schneider et al., 2021). Mayer (2021) provided design principles for embodied instructional videos based on the findings of their research studies. This study went further by conducting a scoping review of the literature on embodied instructional videos and presenting a comprehensive synthesis. The results emphasize different aspects of instructor movements, appearance, and use of instructional media and take into account content elements and learner characteristics, activities, and tools in the design of embodied instructional videos. While this study supports existing embodiment principles to enhance learning, it also reveals the contributing or moderating influence of instructors' deictic gestures, facial expressions, and intimate behaviors, students' prior knowledge and actions during learning, and the complexity of the video subject. Overall, this scoping review advances prior understandings about the embodiment in instructional videos and presents a rich repertoire of design factors. Future research and practices on instructional videos can benefit from the design considerations to produce more effective videos, refine them for specific conditions, and extend them by considering the research gaps revealed in this study.

Supplementary information

Data regarding the studies in the review

ID	Author	Learner group	Location	Video subject	Video type	Instructional media	Video length (avg.)	Treatment duration	Setting	Design factors	Learning process	Learning outcome
1	Alibali et al. (2013)	K-12	USA	Math	Demonstration	Whiteboard	NA	1 session	NA	Gesture (mixed)	NA	L
2	Beege et al. (2017)	Higher education	Germany	Statistics	Lecture	No media	10 min	1 session	Lab	Proximity, orientation	NA	L, SPA
3	Beege et al. (2020), exp. 1 and 2	Higher education	Germany	Geography (exp. 1), history (exp. 2)	Lecture	Pictures	9 min	1 session	Lab	Gesture (beat and deictic)	NA	L, SPA, A, PCP
4	Boucheix et al. (2018)	Higher education	France	Nursing	Demonstration	Instruments	11 min	3 days	NA	Perspective	NA	L, A
5	Carlson et al. (2014)	Higher education	USA	Physics	Lecture	Pictures	NA	1 session	NA	Gesture (iconic), prior knowledge	NA	L
6	Carpenter et al. (2013), exp. 1 and 2	Higher education	USA	Biology	Lecture	No media	1 min	1 session	Lab	Fluency	Behavior (exp. 2)	L, A, SPA
7	Fiorella and Mayer (2016), exp. 1, 2, 3, and 4	Higher education	USA	Physics	Demonstration	Whiteboard	1 min	1 session	Lab	Prior knowledge, gesture (deictic) (exp.1), drawing (exp. 1, 2, and 3), body part (exp. 4)	NA	L
8	Fiorella et al. (2017), exp. 1 and 2	Higher education	USA	Physics	Demonstration	Instruments	1 min	1 session	Lab	Perspective, learning by enactment (exp.1), complexity (exp.1), learning by explaining (exp. 2)	NA	L, PCP
9	Fiorella et al. (2019), exp. 1 and 2	Higher education	USA	Biology	Lecture	Whiteboard	12 min	1 session	Lab	Drawing (exp.1), whiteboard (exp. 2)	NA	L, PCP, SPA, A
10	Ge et al. (2021)	Higher education	China	Foreign language	Lecture	Slides, animation	22 min	1 session	Multimedia classroom	Instructor presence (disregarded), scene image	NA	L, SPA, A
11	Gluhareva and Prieto (2017)	Higher education	Spain	Foreign language	Lecture	No media	7 min	1 session	Lab	Gesture (beat), complexity	NA	L

12	Hekele et al. (2021)	Higher education	Germany	Mechanical/ Material engineering	Demonstration	Instruments, Virtual reality (VR) headset	9 min	1 session	Lab	VR	Eye movement	L, PCP
13	Hirata et al. (2014)	Higher education	USA	Foreign language	Lecture	No media	NA	2 days	NA	Gesture, speaking rate	NA	L
14	Jackson et al. (2006)	Adults	USA	Physical exercise	Demonstration	No media	5 sec	1 session	Lab	Learning by enactment, perspective	Brain activity, behavior	NA
15	Katsioloudis et al. (2013)	Higher education	USA	Mechanical/ Material engineering	Demonstration	Instruments	NA	1 session	NA	Perspective	NA	L
16	Kerssen-Griep and Witt (2015)	Higher education	USA	Communication	Feedback	No media	3-4 min	1 session	Lab	Immediacy, face-threat mitigation	NA	SPA
17	Kizilcec et al. (2015), exp. 2	Adults	NA	Sociology	Lecture	NA	NA	One semester	Online	Duration, learning preference	Behavior	L, SPA, PCP
18	Koumoutsakis et al. (2016)	K-12	USA	Math	Demonstration	Whiteboard	NA	1 session	Mixed	Gesture (deictic)	NA	L
19	Lawson et al. (2021), exp. 1 and 2	Higher education	USA	Statistics	Lecture	Slides	10 min	1 session	Lab	Positivity	NA	L, PCP, A, SPA
20	Londe and Cziraky (2009)	Higher education	USA	Foreign language	Lecture	No media	10 min	1 session	Mixed	Body part	NA	L
21	Lyons (2012)	Higher education	USA	Psychology	Lecture	NA	NA	One semester	Online	Instructor presence (disregarded), technological efficacy	NA	L, SPA, A
22	Merkt et al. (2019), exp. 1 and 2	Higher education	Germany	Biology	Lecture, demonstration	Flipchart, models	13 min	1 session	Lab	Scene image	NA	L, PCP, A, SPA
23	Ouwehand et al. (2015a)	Higher education	Netherlands	Math	Demonstration	Slides	2 min	1 session	Lab	Gaze, gesture (deictic)	Eye movement	L, PCP
24	Ouwehand et al. (2015b)	K12, higher education, older adults	Netherlands	Math	Demonstration	Slides	NA	1 session	Lab	Gesture (deictic), age	NA	L, PCP
25	Perez-Navarro et al. (2021)	Higher education	Spain	Physics	Lecture, demonstration	Whiteboard	NA	One semester	Online	Body part	NA	A
26	Pi et al. (2017a)	Higher education	China	Educational technology	Lecture	Slides	7 min	1 session	Lab	Gesture (deictic)	Eye movement	L, PCP
27	Pi et al. (2017b)	Higher education	China	Educational technology	Lecture	Slides	7 min	1 session	Lab	Size	NA	L, PCP, SPA, A
28	Pi et al. (2019a)	Higher education	China	Educational technology	Lecture	Slides	7 min	1 session	Lab	Gesture (deictic, iconic, and metaphoric), prior knowledge	Eye movement	L

29	Pi et al. (2019b)	Higher education	China	Biology	Lecture	Slides	8 min	1 session	Lab	Gaze, gesture (deictic)	Eye movement	L
30	Pi et al. (2020a)	Higher education	China	Geography	Lecture	Slides	7 min	1 session	Lab	Gaze, facial expression	Eye movement	L
31	Pi et al. (2020b)	Higher education	China	Geography	Lecture	Slides	7 min	1 session	Lab	Gaze, orientation	Eye movement	L
32	Pi et al. (2021a), exp. 1 and 2	Higher education	China	Foreign language	Lecture	Slides	2-5 min (exp. 1), 13 min (exp. 2)	1 session	Lab	Gesture (beat, deictic, and iconic), complexity (exp. 2)	Brain activity (exp. 1)	L
33	Pi et al. (2021c)	Higher education	China	Geography	Lecture	Slides	8 min	1 session	Lab	Gaze, facial expression	Eye movement	L, SPA, PCP
34	Pi et al. (2021d), exp. 1 and 2	Higher education	China	Biology	Lecture	Slides	7 min	1 session	Lab	Gesture (beat), complexity (exp. 2)	Eye movement (exp. 1)	L, PCP, SPA, A
35	Pi et al. (2021e), exp. 1 and 2	Higher education	China	Foreign language	Lecture	Slides	2-5 min	1 session	Lab	Gesture (beat)	Brain activity (exp. 2)	L (exp. 1), PCP (exp. 2)
36	Pilegard and Fiorella (2021), exp. 1 and 2	Higher education	USA	History (exp. 1), biology (exp. 2)	Lecture	No media	1 min	1 session	Lab	Gesture (metaphoric, iconic, and beat)	NA	L, SPA, A
37	Schneider et al. (2021)	Higher education	Germany	Geography	Lecture	Pictures	1 min	1 session	Lab	Gesture (metaphoric), facial expression	NA	L, PCP, SPA
38	Stull et al. (2018a), exp. 1 and 2	Higher education	USA	Chemistry	Demonstration	Whiteboard	20 min	1 session	Classroom	Whiteboard	NA	L, PCP, SPA, A
39	Stull et al. (2018b)	Higher education	USA	Chemistry	Demonstration	Whiteboard	20 min	1 session	Lab	Whiteboard	Eye movement	L, PCP, SPA, A
40	Stull et al. (2018c), exp. 1 and 2	Higher education	USA	Chemistry	Demonstration	Whiteboard, models	20 min (exp. 1), 50 min (exp. 2)	1 session	Classroom (exp. 1), lecture hall (exp. 2)	Learning by enactment	NA	L, PCP, SPA, A
41	Stull et al. (2021)	Higher education	USA	Biology	Lecture	Whiteboard	14 min	1 session	Lab	Gaze, whiteboard	Eye movement	L, PCP, SPA, A
42	Swenson et al. (2021)	Higher education	USA	Mechanical/ Material engineering	Demonstration	Slides, whiteboard	10 min	1 session	Online	Whiteboard, drawing/ writing	NA	A, PCP
43	Toftness et al. (2018), exp. 1 and 2	Higher education	USA	Communication	Lecture	Slides	31 min	1 session	NA	Fluency	NA	L, SPA, A
44	van Gog et al. (2014)	Higher education	Netherlands	Math	Demonstration	Models	2 min	1 session	Lab	Body part	Eye movement	L
45	van Wermeskerken and van Gog (2017)	Higher education	Netherlands	Chemistry	Demonstration	Models	4 min	1 session	Lab	Gaze, body part	Eye movement	L

46	van Wermeskerken et al. (2018)	K-12	Netherlands	Chemistry	Demonstration	Models	4 min	1 session	Lab	Body part, autism spectrum disorder	Eye movement	L
47	Waller et al. (2021)	Higher education	Canada	Psychology	Demonstration	VR headset	5 min	1 session	Lab	VR	NA	A, PCP
48	Wang et al. (2019a)	Higher education	China	Teaching	Lecture	Slides	14 min	1 session	Lab	Facial expression	Emotion	L, SPA, A, PCP
49	Wang et al. (2019b)	Higher education	China	Educational technology	Lecture, demonstration	Slides, screencast	6 min	1 session	Lab	Gaze, knowledge type	Eye movement	L, SPA
50	Witt and Kerssen-Griep (2011)	Higher education	USA	Communication	Feedback	No media	3-4 min	1 session	Lab	Immediacy, face-threat mitigation	NA	SPA
51	Yeo et al. (2017)	K-12	USA	Math	Demonstration	Slides	20 min	1 session	NA	Gesture (deictic)	NA	L
52	Yi et al. (2019)	Higher education	China	Educational technology	Demonstration	Screencast	4 min	1 session	Multimedia classroom	Duration	NA	L, SPA, PCP, A
53	Yu (2021)	Higher education	China	Foreign language	Lecture	Slides	NA	One semester	Online	Duration	NA	L, PCP
54	Yuan et al. (2021), exp. 1	Higher education	China	History	Lecture	Slides	9 min	1 session	Online	Facial attractiveness	NA	L, SPA, A
55	Zhang et al. (2021)	Higher education	China	Psychology	Lecture	Slides	5 min	1 session	Lab	Position	Eye movement	L, A

Notes. Exp. = Experiment, NA = Not available, L = Learning, SPA = Social presence/agency, A = Affect (motivation/satisfaction/emotion), PCP = Perception of cognitive process.

Abbreviations

EEG: Electroencephalography; fMRI: Functional magnetic resonance imaging; MOOC: Massive open online course; STEM: Science, technology, engineering, and math; VR: Virtual reality.

Author's contributions

The author is responsible for the whole manuscript.

Author's information

Ecenaz Alemdag received her Ph.D. degree in Computer Education and Instructional Technology from Middle East Technical University in Ankara, Turkey. Her research interests are feedback, instructional videos, multimedia learning, online learning, and user experience.

Funding

Not applicable.

Availability of data and materials

The dataset supporting the conclusions of this article is presented in Supplementary information.

Declarations

Competing interests

The author declares that she has no competing interests.

Received: 23 May 2022 Accepted: 23 November 2022

Published: 28 February 2023 (Online First: 28 December 2022)

References

*Sample article involved in the review

Akçayır, G., & Akçayır, M. (2018). The flipped classroom: A review of its advantages and challenges. *Computers & Education, 126*, 334–345. https://doi.org/10.1016/j.compedu.2018.07.021

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

Alemdag, E., & Cagiltay, K. (2018). A systematic review of eye tracking research on multimedia learning. *Computers & Education, 125*, 413–428. https://doi.org/10.1016/j.compedu.2018.06.023

Alibali, M. W., & Nathan, M. J. (2012). Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures. *Journal of the Learning Sciences, 21*(2), 247–286. https://doi.org/10.1080/10508406.2011.611446

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology, 8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Atkinson, R. K., Mayer, R. E., & Merrill, M. M. (2005). Fostering social agency in multimedia learning: Examining the impact of an animated agent's voice. *Contemporary Educational Psychology, 30*(1), 117–139. https://doi.org/10.1016/j.cedpsych.2004.07.001

*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 & Education*, *156*, Article 103955. https://doi.org/10.1016/j.compedu.2020.103955

Belt, E. S., & Lowenthal, P. R. (2021). Video use in online and blended courses: A qualitative synthesis. *Distance Education, 42*(3), 410–440. https://doi.org/10.1080/01587919.2021.1954882

Biard, N., Cojean, S., & Jamet, E. (2018). Effects of segmentation and pacing on procedural learning by video. *Computers in Human Behavior, 89*, 411–417. https://doi.org/10.1016/j.chb.2017.12.002

Birgili, B., Seggie, F. N., & Oğuz, E. (2021). The trends and outcomes of flipped learning research between 2012 and 2018: A descriptive content analysis. *Journal of Computers in Education, 8*, 365–394. https://doi.org/10.1007/s40692-021-00183-y

*Boucheix, J. M., Gauthier, P., Fontaine, J. B., & Jaffeux, S. (2018). Mixed camera viewpoints improve learning medical hand procedure from video in nurse training?. *Computers in Human Behavior*, *89*, 418–429. https://doi.org/10.1016/j.chb.2018.01.017

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

*Carlson, C., Jacobs, S. A., Perry, M., & Church, R. B. (2014). The effect of gestured instruction on the learning of physical causality problems. *Gesture*, *14*(1), 26–45. https://doi.org/10.1075/gest.14.1.02car

*Carpenter, S. K., Wilford, M. M., Kornell, N., & Mullaney, K. M. (2013). Appearances can be deceiving: Instructor fluency increases perceptions of learning without increasing actual learning. *Psychonomic Bulletin & Review*, *20*(6), 1350–1356. https://doi.org/10.3758/s13423-013-0442-z

Carroll, J. M., & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology, 70*(2), 205–218. https://doi.org/10.1037/0022-3514.70.2.205

Castro-Alonso, J. C., Wong, R. M., Adesope, O. O., & Paas, F. (2021). Effectiveness of multimedia pedagogical agents predicted by diverse theories: A meta-analysis. *Educational Psychology Review, 33*(3), 989–1015. https://doi.org/10.1007/s10648-020-09587-1

Chorianopoulos, K. (2018). A taxonomy of asynchronous instructional video styles. *International Review of Research in Open and Distributed Learning, 19*(1), 294–311. https://doi.org/10.19173/irrodl.v19i1.2920

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. (2001). What is next in the media and methods debate?. In R.E. Clark (Ed.), *Learning from media arguments, analysis, and evidence* (pp. 327–337). Information Age Publishing.

Consolvo, S., Bentley, F. R., Hekler, E. B., & Phatak, S. S. (2017). *Mobile user research: A practical guide*. Morgan & Claypool Publishers.

Craig, S. D., Twyford, J., Irigoyen, N., & Zipp, S. A. (2015). A test of spatial contiguity for virtual human's gestures in multimedia learning environments. *Journal of Educational Computing Research, 53*(1), 3–14. https://doi.org/10.1177/0735633115585927

Davis, R. O. (2018). The impact of pedagogical agent gesturing in multimedia learning environments: A meta-analysis. *Educational Research Review, 24*, 193–209. https://doi.org/10.1016/j.edurev.2018.05.002

de Koning, B. B., Hoogerheide, V., & Boucheix, J. M. (2018). Developments and trends in learning with instructional video. *Computers in Human Behavior, 89*, 395–398. https://doi.org/10.1016/j.chb.2018.08.055

Denisova, A., & Cairns, P. (2015). First person vs. third person perspective in digital games: Do player preferences affect immersion?. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 145–148). Association for Computing Machinery.

Doherty, C. (2022). An investigation into the relationship between multimedia lecture design and learners' engagement behaviours using web log analysis. *Plos One, 17*(8), Article e0273007. https://doi.org/10.1371/journal.pone.0273007

Fidan, M., & Debbag, M. (2022). Comparing the effectiveness of instructional video types: An in-depth analysis on pre-service teachers for online learning. *International Journal of Human–Computer Interaction*, 1–12. https://doi.org/10.1080/10447318.2022.2041905

Fiorella, L., & Mayer, R. E. (2015). *Learning as a generative activity: Eight learning strategies that promote understanding*. Cambridge University Press.

*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., & Mayer, R. E. (2018). What works and doesn't work with instructional video. *Computers in Human Behavior, 89*, 465–470. https://doi.org/10.1016/j.chb.2018.07.015

*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

Fiorella, L., Stull, A. T., Kuhlmann, S., & Mayer, R. E. (2020). Fostering generative learning from video lessons: Benefits of instructor-generated drawings and learner-generated explanations. *Journal of Educational Psychology, 112*(5), 895–906. http://dx.doi.org/10.1037/edu0000408

*Fiorella, L., van Gog, T., Hoogerheide, V., & Mayer, R. E. (2017). It's all a matter of perspective: Viewing first-person video modeling examples promotes learning of an assembly task. *Journal of Educational Psychology*, *109*(5), 653–665. https://doi.org/10.1037/edu0000161

Frankel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education* (8th ed.). The McGraw-Hill Companies, Inc.

Fyfield, M. (2022). YouTube in the secondary classroom: How teachers use instructional videos in mainstream classrooms. *Technology, Pedagogy and Education, 31*(2), 185–197. https://doi.org/10.1080/1475939X.2021.1980429

Fyfield, M., Henderson, M., & Phillips, M. (2022). Improving instructional video design: A systematic review. *Australasian Journal of Educational Technology, 38*(3), 150–178. https://doi.org/10.14742/ajet.7296

Fyfield, M., Henderson, M., Heinrich, E., & Redmond, P. (2019). Videos in higher education: Making the most of a good thing. *Australasian Journal of Educational Technology, 35*(5), 1–7. https://doi.org/10.14742/ajet.5930

*Gluhareva, D., & Prieto, P. (2017). Training with rhythmic beat gestures benefits L2 pronunciation in discourse-demanding situations. *Language Teaching Research*, *21*(5), 609–631. https://doi.org/10.1177/1362168816651463

Gorisse, G., Christmann, O., Amato, E. A., & Richir, S. (2017). First-and third-person perspectives in immersive virtual environments: presence and performance analysis of embodied users. *Frontiers in Robotics and AI, 4*, Article 33. https://doi.org/10.3389/frobt.2017.00033

Gullberg, M. (2006). Some reasons for studying gesture and second language acquisition (Hommage à Adam Kendon). *IRAL: International Review of Applied Linguistics in Language Teaching, 44*, 103–124.

Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: An empirical study of MOOC videos. In *Proceedings of the First ACM Conference on Learning @ Scale Conference* (pp. 41–50). Association for Computing Machinery.

Harrison, T. (2020). How distance education students perceive the impact of teaching videos on their learning. *Open Learning: The Journal of Open, Distance and e-Learning, 35*(3), 260–276. https://doi.org/10.1080/02680513.2019.1702518

Heidig, S., & Clarebout, G. (2011). Do pedagogical agents make a difference to student motivation and learning?. *Educational Research Review, 6*(1), 27–54. https://doi.org/10.1016/j.edurev.2010.07.004

*Hekele, F., Spilski, J., Bender, S., & Lachmann, T. (2021). Remote vocational learning opportunities-A comparative eye-tracking investigation of educational 2D videos versus 360° videos for car mechanics. *British Journal of Educational Technology*, 1–21. https://doi.org/10.1111/bjet.13162

Henderson, M., & Phillips, M. (2015). Video-based feedback on student assessment: Scarily personal. *Australasian Journal of Educational Technology, 31*(1), 51–66.

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*, Article 100059. https://doi.org/10.1016/j.caeo.2021.100059

Hung, S. T. A. (2016). Enhancing feedback provision through multimodal video technology. *Computers & Education, 98*, 90–101. https://doi.org/10.1016/j.compedu.2016.03.009

Ibrahim, M., Antonenko, P. D., Greenwood, C. M., & Wheeler, D. (2012). Effects of segmenting, signalling, and weeding on learning from educational video. *Learning, Media and Technology, 37*(3), 220–235. https://doi.org/10.1080/17439884.2011.585993

*Jackson, P. L., Meltzoff, A. N., & Decety, J. (2006). Neural circuits involved in imitation and perspective-taking. *Neuroimage*, *31*(1), 429–439. https://doi.org/10.1016/j.neuroimage.2005.11.026

Jewitt, C., Bezemer, J., & O'Halloran, K. (2016). *Introducing multimodality*. Routledge.

Kalyuga, S. (2014). The expertise reversal principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 576–597). Cambridge University Press.

*Kerssen-Griep, J., & Witt, P. L. (2015). Instructional feedback III: How do instructor facework tactics and immediacy cues interact to predict student perceptions of being mentored?. *Communication Education*, *64*(1), 1–24. https://doi.org/10.1080/03634523.2014.978797

*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

Kokoç, M. (2019). Effects of using instructor presence in programming video lectures: An eye tracking analysis. *Ege Journal of Educational Technologies, 3*(2), 52–66.

*Koumoutsakis, T., Church, R. B., Alibali, M. W., Singer, M., & Ayman-Nolley, S. (2016). Gesture in instruction: Evidence from live and video lessons. *Journal of Nonverbal Behavior*, *40*(4), 301–315. https://doi.org/10.1007/s10919-016-0234-z

Kruger, J. L., & Doherty, S. (2016). Measuring cognitive load in the presence of educational video: Towards a multimodal methodology. *Australasian Journal of Educational Technology, 32*(6), 19–31. http://dx.doi.org/10.14742/ajet.3084

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. https://doi.org/10.2307/2529310

Lange, C., & Costley, J. (2020). Improving online video lectures: Learning challenges created by media. *International Journal of Educational Technology in Higher Education, 17*(1), 1–18. https://doi.org/10.1186/s41239-020-00190-6

*Lawson, A. P., Mayer, R. E., Adamo-Villani, N., Benes, B., Lei, X., & Cheng, J. (2021). The positivity principle: Do positive instructors improve learning from video lectures?. *Educational Technology Research and Development*, *69*(6), 3101–3129. https://doi.org/10.1007/s11423-021-10057-w

Li, W., Wang, F., Mayer, R. E., & Liu, H. (2019). Getting the point: Which kinds of gestures by pedagogical agents improve multimedia learning?. *Journal of Educational Psychology, 111*(8), 1382–1395. https://doi.org/10.1037/edu0000352

Lim, F. V. (2021). *Designing learning with embodied teaching: Perspectives from multimodality*. Routledge.

Lodico, M. G., Spaulding, D. T., & Voegtle, K. H. (2010). *Methods in educational research: From theory to practice* (2nd ed.). John Wiley & Sons.

Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research, 28*(4), 587–604. https://doi.org/10.1111/j.1468-2958.2002.tb00826.x

Lusk, M. M., & Atkinson, R. K. (2007). Animated pedagogical agents: Does their degree of embodiment impact learning from static or animated worked examples?. *Applied Cognitive Psychology, 21*(6), 747–764. https://doi.org/10.1002/acp.1347

*Lyons, A., Reysen, S., & Pierce, L. (2012). Video lecture format, student technological efficacy, and social presence in online courses. *Computers in Human Behavior*, *28*(1), 181–186. https://doi.org/10.1016/j.chb.2011.08.025

Mayer, R. E. (2014a). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 43–71). Cambridge University Press.

Mayer, R. E. (2014b). Principles based on social cues in multimedia learning: Personalisation, voice, image, and embodiment principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 345–368). Cambridge University Press.

Mayer, R. E. (2021). Evidence-based principles for how to design effective instructional videos. *Journal of Applied Research in Memory and Cognition, 10*(2), 229–240. https://doi.org/10.1016/j.jarmac.2021.03.007

Mayer, R. E., & DaPra, C. S. (2012). An embodiment effect in computer-based learning with animated pedagogical agents. *Journal of Experimental Psychology: Applied, 18*(3), 239–252. https://doi.org/10.1037/a0028616

Mayer, R. E., Fiorella, L., & Stull, A. (2020). Five ways to increase the effectiveness of instructional video. *Educational Technology Research and Development, 68*(3), 837–852. https://doi.org/10.1007/s11423-020-09749-6

Mayer, R. E., Sobko, K., & Mautone, P. D. (2003). Social cues in multimedia learning: Role of speaker's voice. *Journal of Educational Psychology, 95*(2), 419–425. https://doi.org/10.1037/0022-0663.95.2.419

Mays, N., Roberts, E., & Popay, J. (2001). Synthesising research evidence. In N. Fulop, P. Allen, A. Clarke & N. Black (Eds.), *Studying the organisation and delivery of health services: Research methods* (pp. 188–220). Routledge.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.

Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & Education, 70*, 29–40. https://doi.org/10.1016/j.compedu.2013.07.033

*Merkt, M., Lux, S., Hoogerheide, V., van Gog, T., & Schwan, S. (2020). A change of scenery: Does the setting of an instructional video affect learning? *Journal of Educational Psychology, 112*(6), 1273–1283. https://doi.org/10.1037/edu0000414

Moon, J., & Ryu, J. (2021). 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*(1), 39–63. https://doi.org/10.1007/s12528-020-09255-x

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review, 19*(3), 309–326. https://doi.org/10.1007/s10648-007-9047-2

Munn, Z., Peters, M. D., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology, 18*(1), 1–7. https://doi.org/10.1186/s12874-018-0611-x

Mutlu-Bayraktar, D., Cosgun, V., & Altan, T. (2019). Cognitive load in multimedia learning environments: A systematic review. *Computers & Education, 141*, Article 103618. https://doi.org/10.1016/j.compedu.2019.103618

Ou, C., Joyner, D. A., & Goel, A. K. (2019). Designing and developing video lessons for online learning: A seven-principle model. *Online Learning, 23*(2), 82–104. https://doi.org/10.24059/olj.v23i2.1449

*Ouwehand, K., van Gog, T., & Paas, F. (2015a). Designing effective video-based modeling examples using gaze and gesture cues. *Journal of Educational Technology & Society*, *18*(4), 78–88.

*Ouwehand, K., van Gog, T., & Paas, F. (2015b). Effects of gestures on older adults' learning from video-based models. *Applied Cognitive Psychology*, *29*(1), 115–128. https://doi.org/10.1002/acp.3097

Ozel, P., Mutlu-Bayraktar, D., Altan, T., Coskun, V., & Olamat, A. (2021). Neuroimaging tools in multimedia learning: A systematic review. *Interactive Learning Environments*, 1–18. https://doi.org/10.1080/10494820.2021.1984255

Pass, F, & Sweller, J. (2014). Implications of cognitive load for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 27–42). Cambridge University Press.

Peters, M. D., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *JBI Evidence Implementation, 13*(3), 141–146. https://doi.org/10.1097/XEB.0000000000000050

*Pi, Z., Hong, J., & Yang, J. (2017a). Effects of the instructor's pointing gestures on learning performance in video lectures. *British Journal of Educational Technology*, *48*(4), 1020–1029. https://doi.org/10.1111/bjet.12471

*Pi, Z., Hong, J., & Yang, J. (2017b). Does instructor's image size in video lectures affect learning outcomes?. *Journal of Computer Assisted Learning*, *33*(4), 347–354. https://doi.org/10.1111/jcal.12183

*Pi, Z., Zhang, Y., Yang, J., Hu, W., & Yang, H. H. (2019a). All roads lead to Rome: Instructors' pointing and depictive gestures in video lectures promote learning through different patterns of attention allocation. *Journal of Nonverbal Behavior*, *43*(4), 549–559. https://doi.org/10.1007/s10919-019-00310-5

*Pi, Z., Zhang, Y., Yu, Q., Zhang, Y., Yang, J., & Zhao, Q. (2021a). Neural oscillations and learning performance vary with an instructor's gestures and visual materials in video lectures. *British Journal of Educational Technology*, 1–21. https://doi.org/10.1111/bjet.13154

Pi, Z., Zhang, Y., Zhou, W., Xu, K., Chen, Y., Yang, J., & Zhao, Q. (2021b). Learning by explaining to oneself and a peer enhances learners' theta and alpha oscillations while watching video lectures. *British Journal of Educational Technology, 52*(2), 659–679. https://doi.org/10.1111/bjet.13048

*Pi, Z., Zhang, Y., Zhu, F., Chen, L., Guo, X., & Yang, J. (2021c). The mutual influence of an instructor's eye gaze and facial expression in video lectures. *Interactive Learning Environments*, 1–18. https://doi.org/10.1080/10494820.2021.1940213

*Pi, Z., Zhang, Y., Zhu, F., Xu, K., Yang, J., & Hu, W. (2019b). Instructors' pointing gestures improve learning regardless of their use of directed gaze in video lectures. *Computers & Education*, *128*, 345–352. https://doi.org/10.1016/j.compedu.2018.10.006

*Pi, Z., Zhu, F., Zhang, Y., Chen, L., & Yang, J. (2021d). Complexity of visual learning material moderates the effects of instructor's beat gestures and head nods in video lectures. *Learning and Instruction*, *77*, Article 101520. https://doi.org/10.1016/j.learninstruc.2021.101520

*Pi, Z., Zhu, F., Zhang, Y., & Yang, J. (2021e). An instructor's beat gestures facilitate second language vocabulary learning from instructional videos: Behavioral and neural evidence. *Language Teaching Research*, 1–29. https://doi.org/10.1177/13621688211039023

*Pilegard, C., & Fiorella, L. (2021). Using gestures to signal lesson structure and foster meaningful learning. *Applied Cognitive Psychology*, *35*(5), 1362–1369. https://doi.org/10.1002/acp.3866

Polat, H., Kayaduman, H., & Battal, A. (2021). Video ders materyali tasarımında bir ikilem: Öğreticinin görüntüsü [A dilemma in video course material design: Teacher's image]. In H. E. Odabaşı, B. Akkoyunlu & A. İşman (Eds.), *Eğitim Teknolojisi Okumaları 2021 [Educational Technology Readings 2021]* (pp. 575–597). Pegem Akademi.

Rodemer, M., Lindner, M. A., Eckhard, J., Graulich, N., & Bernholt, S. (2022). Dynamic signals in instructional videos support students to navigate through complex representations: An eye-tracking study. *Applied Cognitive Psychology, 36*(4), 852–863. https://doi.org/10.1002/acp.3973

Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010). A closer look at split visual attention in system-and self-paced instruction in multimedia learning. *Learning and Instruction, 20*(2), 100–110. https://doi.org/10.1016/j.learninstruc.2009.02.011

*Schneider, S., Krieglstein, F., Beege, M., & Rey, G. D. (2021). The impact of video lecturers' nonverbal communication on learning–An experiment on gestures and facial expressions of pedagogical agents. *Computers & Education*, *176*, Article 104350. https://doi.org/10.1016/j.compedu.2021.104350

Schroeder, N. L., Adesope, O. O., & Gilbert, R. B. (2013). How effective are pedagogical agents for learning? A meta-analytic review. *Journal of Educational Computing Research, 49*(1), 1–39. https://doi.org/10.2190/EC.49.1.a

*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

*Stull, A. T., Fiorella, L., Gainer, M. J., & Mayer, R. E. (2018a). Using transparent whiteboards to boost learning from online STEM lectures. *Computers & Education*, *120*, 146–159. https://doi.org/10.1016/j.compedu.2018.02.005

*Stull, A. T., Fiorella, L., & Mayer, R. E. (2018b). An eye-tracking analysis of instructor presence in video lectures. *Computers in Human Behavior*, *88*, 263–272. https://doi.org/10.1016/j.chb.2018.07.019

*Stull, A. T., Gainer, M. J., & Hegarty, M. (2018c). Learning by enacting: The role of embodiment in chemistry education. *Learning and Instruction*, *55*, 80–92. https://doi.org/10.1016/j.learninstruc.2017.09.008

*Swenson, M. J., Spence, T., & Smentkowski, B. (2021). Student-led development of a lightboard to enhance future student learning. *International Journal of Mechanical Engineering Education*, 1–16. https://doi.org/10.1177/03064190211026229

*Toftness, A. R., Carpenter, S. K., Geller, J., Lauber, S., Johnson, M., & Armstrong, P. I. (2018). Instructor fluency leads to higher confidence in learning, but not better learning. *Metacognition and Learning*, *13*(1), 1–14. https://doi.org/10.1007/s11409-017-9175-0

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., … Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMAScR): Checklist and explanation. *Annals of Internal Medicine, 169*(7), 467–473. https://doi.org/10.7326/M18-0850

Tseng, S. S. (2021). The influence of teacher annotations on student learning engagement and video watching behaviors. *International Journal of Educational Technology in Higher Education, 18*(1), 1–17. https://doi.org/10.1186/s41239-021-00242-5

*van Gog, T., Verveer, I., & Verveer, L. (2014). Learning from video modeling examples: Effects of seeing the human model's face. *Computers & Education*, *72*, 323–327. https://doi.org/10.1016/j.compedu.2013.12.004

*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 & Education*, *113*, 98–107. https://doi.org/10.1016/j.compedu.2017.05.013

*van Wermeskerken, M., Grimmius, B., & van Gog, T. (2018). Attention to the model's face when learning from video modeling examples in adolescents with and without autism spectrum disorder. *Journal of Computer Assisted Learning*, *34*(1), 32–41. https://doi.org/10.1111/jcal.12211

Violante, M. G., Vezzetti, E., & Piazzolla, P. (2019). Interactive virtual technologies in engineering education: Why not 360° videos?. *International Journal on Interactive Design and Manufacturing (IJIDeM), 13*(2), 729–742. https://doi.org/10.1007/s12008-019-00553-y

*Waller, M., Mistry, D., Jetly, R., & Frewen, P. (2021). Meditating in Virtual Reality 3: 360° Video of Perceptual Presence of Instructor. *Mindfulness*, *12*(6), 1424–1437. https://doi.org/10.1007/s12671-021-01612-w

Wang, F., Li, W., Mayer, R. E., & Liu, H. (2018). Animated pedagogical agents as aids in multimedia learning: Effects on eye-fixations during learning and learning outcomes. *Journal of Educational Psychology, 110*(2), 250–268. http://dx.doi.org/10.1037/edu0000221

*Wang, Y., Liu, Q., Chen, W., Wang, Q., & Stein, D. (2019a). Effects of instructor's facial expressions on students' learning with video lectures. *British Journal of Educational Technology*, *50*(3), 1381–1395. https://doi.org/10.1111/bjet.12633

*Wang, H., Pi, Z., & Hu, W. (2019b). The instructor's gaze guidance in video lectures improves learning. *Journal of Computer Assisted Learning*, *35*(1), 42–50. https://doi.org/10.1111/jcal.12309

Wang, J., Antonenko, P., Keil, A., & Dawson, K. (2020). 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

*Witt, P. L., & Kerssen-Griep, J. (2011). Instructional feedback I: The interaction of facework and immediacy on students' perceptions of instructor credibility. *Communication Education*, *60*(1), 75–94. https://doi.org/10.1080/03634523.2010.507820

Wong, R. M., & Adesope, O. O. (2021). Meta-analysis of emotional designs in multimedia learning: A replication and extension study. *Educational Psychology Review, 33*(2), 357–385. https://doi.org/10.1007/s10648-020-09545-x

Wouters, P., Paas, F., & van Merriënboer, J. J. G. (2008). How to optimize learning from animated models: A review of guidelines based on cognitive load. *Review of Educational Research, 78*(3), 645–675. https://doi.org/10.3102/0034654308320320

*Yeo, A., Ledesma, I., Nathan, M. J., Alibali, M. W., & Church, R. B. (2017). Teachers' gestures and students' learning: Sometimes "hands off" is better. *Cognitive Research: Principles and Implications*, *2*(1), 1–11. https://doi.org/10.1186/s41235-017-0077-0

*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

*Yu, Z. (2021). The effect of teacher presence in videos on intrinsic cognitive loads and academic achievements. *Innovations in Education and Teaching International*, 1–12. https://doi.org/10.1080/14703297.2021.1889394

*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*, 1–14. https://doi.org/10.3389/fpsyg.2021.746857

Publisher's Note