Is student learning from a video lecture affected by whether the instructor wears a mask?

Article in Applied Cognitive Psychology · January 2024

DOI: 10.1002/acp.4169

3 authors:

Xiaoxue Leng

Fuxing Wang
Central China Normal University

Richard Mayer
University of California, Santa Barbara

Received: 24 November 2022 | Revised: 10 October 2023 | Accepted: 22 December 2023

DOI: 10.1002/acp.4169

RESEARCH ARTICLE

WILEY

Is student learning from a video lecture affected by whether the instructor wears a mask?

Xiaoxue Leng[1] | Fuxing Wang[1] | Richard E. Mayer[2]

[1]School of Psychology, Central China Normal University, Wuhan, China

[2]Department of Psychological and Brain Sciences, University of California, Santa Barbara, Santa Barbara, California, USA

Correspondence
Fuxing Wang, School of Psychology, Central China Normal University, 382 Xiongchu Ave, Nanhu BLDG 8073, Wuhan 430079, China.
Email: fxwang@ccnu.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Number: 62277025

Abstract

This study examined whether having the instructor wear a mask during a video lecture affects learning. In Experiment 1, college students watched an instructional video on the formation of lightning, in which an instructor who either did or did not wear a mask as she stood next to slides and lectured. Learners' learning outcomes did not differ significantly, but learners spent significantly less time looking at the instructor's face when she was masked. In Experiment 2, using a 2 (the instructor wore a mask or not) $\times$ 2 (slides were displayed or not) between-subject design, college students learned about the process of water cycle from instructional videos. There was a significant interaction in which adding slides improved learning outcomes with a masked instructor, but not with an unmasked instructor. Adding a mask lowered student ratings of social presence with the instructor. Practical and theoretical implications are discussed.

KEYWORDS

instructional video, learning with video, mask, video lectures, visual attention

1 | INTRODUCTION

1.1 | Objective and rationale

Consider a scenario in which a learner views an instructional video based on a classroom lecture. The instructor stands next to a series of slides as she explains a scientific topic such as how lightning storms develop or how the water cycle works, such as exemplified in the top panel of Figure 1. Would it matter if the instructor wore a mask, as is typical during the pandemic, such as exemplified in the bottom panel of Figure 1? This is both a practical and theoretical question concerning the design of instructional video (Fiorella, 2022; Mayer et al., 2020).

During the COVID-19 pandemic, wearing masks became a recommended method to protect people from the virus, including in college classrooms. However, except for decreasing the spread of virus, there are several potential concerns about wearing masks in educational settings. When an instructor wears a mask, students cannot see instructor's mouth. As the mouth region is covered, students are not able to lipread, which might damage their comprehension of the learning content (e.g., Jordan & Thomas, 2011). Furthermore, masking of instructors makes it more difficult for learners to receive social cues (e.g., Marini et al., 2021), which might reduce student motivation and learning performance. Researchers have raised these concerns about the impact of masks in educational settings (see Spitzer, 2020, for a review). However, there is little research to examine the impact of wearing masks in educational settings, and the effect of masks on students' learning outcomes and learning processes is still unclear. The goal of this study was to investigate how the instructor wearing a mask would influence students' visual attention, subjective perception, and learning performance. Considering that it was difficult to conduct a field study during lockdown, in the current study, we investigated these questions in a multimedia learning environment using a recorded video clip. Further, this study can provide practical guidance for video recording in situations where people have to wear masks such as chemical labs or hospitals.

1.2 | Literature review

1.2.1 | Wearing a mask and learning

Teachers' mask wearing might damage learning because it covers the bottom half region of the teacher's face, which could interfere with

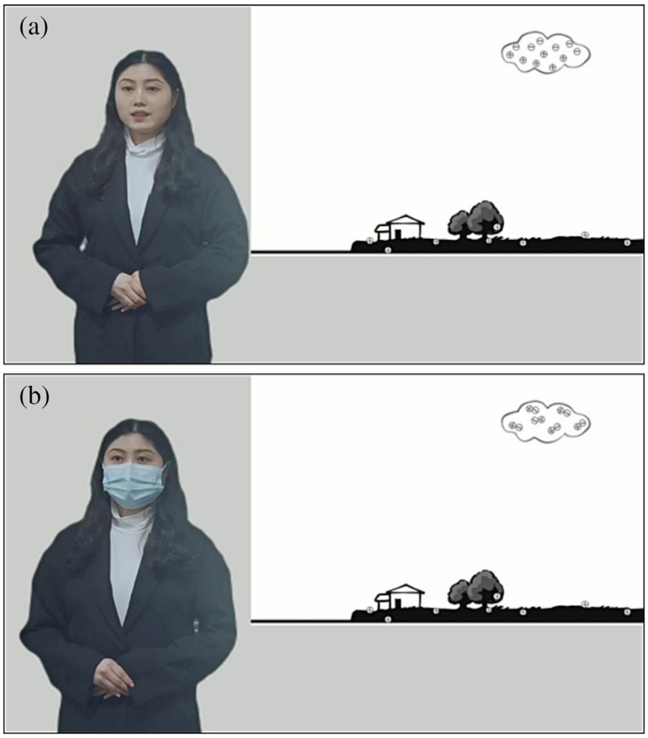

FIGURE 1 Example from the video lessons on the formation of lightning. Panel a: An example of the no-mask group; Panel b: An example of the mask group.

the learner's speech comprehension (e.g., IJsseldijk, 1992; Jordan & Thomas, 2011). Previous studies had demonstrated that when people could see the mouth moving and listen to the speaker simultaneously, their speech perception was better than when listening only (Calvert et al., 1997; Erber, 1975; Sommers & Phelps, 2016). Davis and Kim (2006) provided empirical evidence that, compared with watching the whole face, participants recognized fewer words correctly when only seeing the upper half of the face. Similarly, Jordan and Thomas (2011) found participants made more errors in speech perception when watching speakers who had their lower face occluded. Eye movement studies also indicated the importance of the bottom half of face during speech perception, in that observers allocated much of their visual attention focusing on the bottom half of the speaker's face in speech perception (e.g., Buchan et al., 2007, 2008; Lansing & McConkie, 1999, 2003; Lusk & Mitchel, 2016; Rennig & Beauchamp, 2018). For example, Lansing and McConkie (1999) reported that learners allocated more than 70% of their gaze duration on the bottom of the face (including nose, mouth, and cheek) when watching speakers. In a similar study, Lusk and Mitchel (2016) reported that learners allocated around 274 s total gaze duration on speaker's mouth and nose, and about 163 s on the speaker's eyes.

Moreover, neuroimaging studies also supported the idea that non-verbal cues provided by mouth movements facilitated speech perception (e.g., Giordano et al., 2017; Park et al., 2016; Zhang & Du, 2022). For example, Haider et al. (2022) showed that covering a speaker's lower face with a face mask impaired the learner's neural tracking of speech features. This work suggests that when the speaker wears a mask, cognitive cues provided by the bottom half of the face region are covered and the listener is not able to lipread through masks, so their learning might be impaired.

The bottom half of the face also offers social cues, which can influence students' motivation and learning outcomes. In the research field of pedagogical agents (i.e., onscreen characters or human), social agency theory and the embodiment principle suggest that social cues provided by agents promote students' learning. Specifically, the embodiment principle posits that humanlike onscreen agents help learners learn more deeply and perform better on transfer tests (Fiorella & Mayer, 2022; Mayer, 2021). According to social agency theory, social cues provided by pedagogical agents such as gestures, eye gaze, facial expression, and other body movements help trigger a social response in learners, resulting in learners treating pedagogical agents as their social partners. When the feeling of social presence is increased, pedagogical agents are more likely to be perceived as real people. This primes rules of social communication in which learners exert more effort to select relevant information, organize information into representations, and integrate multiple representations into one coherent representation. Finally, learners engage in better learning processing resulting in better posttest performance (Fiorella & Mayer, 2022; Mayer, 2021; Mayer & DaPra, 2012; Wang et al., 2018, 2022).

Some studies provided empirical evidence and suggested that, compared to low-embodied agents, high-embodied agents who displayed more human-like behaviors helped students deeply learn more (Li et al., 2019; Mayer, 2021; Mayer & DaPra, 2012; Wang et al., 2018). Research has found that face masks impair social communication, and reduce people's facial perception (Marini et al., 2021; McCrackin et al., 2023; Rinck et al., 2022). Thus, we hypothesized that compared to an instructor who wears a mask, an instructor who does not wear a mask provides better social cues that make the instructor seem more humanlike, therefore, causing better learning outcomes. In conclusion, we hypothesized that learners would learn better from instructors that are not masked than those who are masked, as indicated by posttest scores.

1.2.2 | Wearing a mask and visual attention

People's attention tends to be drawn automatically to the speaker's face (e.g., Gullberg & Holmqvist, 2006). Dynamic mouth movements and the instructor's face might distract the learners' attention, resulting in split attention between the instructor's face and the slides (Ayres & Sweller, 2021; van Wermeskerken & van Gog, 2017). Thus, we hypothesize that when the instructor's mask covers the bottom half of their face, the students' visual attention would focus more on the slide rather than the face.

It should be noted that the split attention effect between attending to the learning content and attending to the instructor's face is not always detrimental in the field of multimedia learning. For example, Kizilcec et al. (2014) found split attention between content and instructor's face, but learning outcomes remained unaffected. van Wermeskerken et al. (2018) found similar results. Some studies even showed that instructor presence on the screen improved learning when attention was drawn by the instructor (Wang & Antonenko, 2017; Wang et al., 2020). Only a few studies found that adding a talking head impaired learning (Sondermann & Merkt, 2023).

Therefore, based on split attention effect, we hypothesized students' attention would be distracted by an unmasked face when slides are displayed; that is, students would fixate more on the bottom half of the face region when the instructor shows their whole face and fixate less on the face region and more on the slides when the instructor wears a mask. Eye-tracking technology provides insight into students' visual attention in the learning process, and it has been widely used in the research of learning in multimedia environments (Hyönä, 2010; Mayer, 2010; Pi et al., 2019; Stull et al., 2018; Wang et al., 2018), so we used eye-tracking technology to address this hypothesis.

When attention is divided, extraneous cognitive load (i.e., cognitive processing that does not support the learning objective and is imposed by poor instructional presentation; Mayer, 2021; Sweller et al., 2011) might be increased. According to the split attention principle (Ayres & Sweller, 2021), extraneous cognitive load will be created when learners seek to attend to disparate visual areas of the screen at the same time. When attention is drawn by the face, learners might experience higher extraneous load. In contrast, when face is covered, learners may mainly focus on the learning content, so that they experience less extraneous load. Thus, we hypothesized wearing a mask would reduce extraneous cognitive load based on the split attention principle. However, there was no clear empirical evidence for the impact of pedagogical agents on extraneous cognitive load (Schroeder & Adesope, 2014). In addition, Colliot and Jamet (2018) found instructor's presence did not influence extraneous cognitive load even when the instructor drew the learners' attention. Therefore, a null effect for extraneous cognitive load is possible in the present study.

1.2.3 | Displaying slides moderated the impact of wearing a mask

In real classroom environments, instructors often use slides showing significant information, keywords, pictures and relevant illustrations to help students process the knowledge. Adding illustrations has been proven to be effective in multimedia learning environment (see Mayer, 2022). Pi and Hong's research (2016) showed that adding slides significantly promoted learning performance. Therefore, we hypothesized adding slides would improve learning outcomes and decrease intrinsic cognitive load.

When only the instructor is presented with no concurrent slides, there is only the effect of wearing a mask. Face masks may occlude cognitive and social cues provided by the bottom half of face (e.g., IJsseldijk, 1992; Thomas & Jordan, 2004), and thus, impair learning. When slides and the instructor are shown simultaneously in a video lecture, students have to distribute their attention to the two elements (e.g., Kizilcec et al., 2014; Wang & Antonenko, 2017). Under such circumstances, the presence of instructors on the screen can cause split attention, which might have negative impacts. However, the split attention effect is not always detrimental to learning outcomes (e.g., Kizilcec et al., 2014). If split attention impairs learning, attention drawn by the instructor's face will lower learning outcome performance; that is, the negative impacts caused by face mask in the slides condition will be less than in the no slides condition. If split attention is not damaging, the negative impacts of face masks can be equivalent. Therefore, there is also a need to investigate whether instructor's wearing a mask influenced students' learning differently when displaying slides or not.

1.3 | Hypotheses and research questions

In the current study, we examined whether the instructor wearing a mask would influence learners' learning outcomes, visual attention, and learner perceptions.

In Experiment 1, we compared visual attention and learning outcomes from a video lesson containing a masked or unmasked instructor standing next to a series of slides as she lectured. Based on the foregoing analyses, our primary prediction is that face masks would impair learning outcomes (hypothesis 1). Based on the split-attention hypothesis, we predicted that students are less likely to look at the instructor and more likely to look at slides when the instructor is masked (hypothesis 2). Finally, face masks reduce the richness of the instructor, which should lead to a reduction of extraneous load (hypothesis 3).

In Experiment 2, we also added conditions in which a masked or unmasked instructor lectured without any slides and we added ratings of social presence and affect as new dependent measures. As in Experiment 1, we expected masking to lead to lower performance on tests of learning outcome (hypothesis 1) and less extraneous load (hypothesis 3) when slides were also present. According to social agency theory, students should give lower ratings of social presence with the instructor when she is masked than when she is unmasked (hypothesis 4). Furthermore, Heidig and Clarebout (2011) suggested that high-embodied agents might provide positive affective and motivational cues, so we proposed there would be a lower level of self-reported affective evaluation when the instructor was wearing a mask that blocked some of those cues (hypothesis 5). Adding slides would improve learning outcomes and decrease intrinsic cognitive load (hypothesis 6). In addition, we investigated whether there is an interaction effect between mask and slides. As stated before, if split attention effect was detrimental, attention drawn by the instructor's face would lower learning performance; that is, the negative impacts caused by face mask in the slides condition will be less than in the no slides condition. Therefore, we hypothesized negative impacts of masking would be less in the slide conditions than in the no slides conditions (hypothesis 7).

2 | EXPERIMENT 1

2.1 | Method

2.1.1 | Participants and design

A prior analysis of G*Power 3.1 suggested a minimum of 42 participants when using a large effect size $d = 0.80$ (which we expected based on the work of Jordan & Thomas, 2011) with power set to 0.80 and alpha set to .05 (Faul et al., 2007). Accordingly, sixty-one participants were recruited from Central China Normal University, all of whom had normal or corrected-to-normal visions. The mean age was 21.31 years (SD = 1.58), and there were 49 women and 12 men. The study was approved by the ethics committee of the university.

The experiment used a one-factor between-subject design. Participants were randomly assigned to two groups: 30 in the mask group and 31 in the no-mask group. There was no significant difference in prior knowledge ($t(59) = -0.10$, $p > .05$), mean age (mask group: $M = 20.13$, SD = 1.41; no-mask group: $M = 20.48$, SD = 1.73; $t(59) = -0.87$, $p > .05$), and proportion of women and men ($\chi^2(1) = 0.34$, $p > .05$) between the two groups.

Because of poor eye-tracking data quality (sampling rate less than 75%), 5 participants were excluded from the analysis of eye-tracking data. Eye-tracking data of 56 participants were analyzed (mask group: 27; no-mask group: 29). The mean age was 20.21 years (SD = 1.54) and there were 11 men and 45 women. There was no significant difference on prior knowledge, $t(55) = 0.04$, $p > .05$, mean age (mask group: $M = 20.11$, SD = 1.45; no-mask group: $M = 20.31$, SD = 1.65; $t(55) = -0.49$, $p > .05$), and proportion of women and men ($\chi^2(1) = 0.89$, $p > .05$) between the two groups.

2.1.2 | Materials

The instructional materials, translated and adapted from Moreno and Mayer (1999), consisted of two versions of a continuous narrated animation that explained major steps in the formation of lightning. In one version the instructor wore a mask (mask condition) and in the other the instructor did not wear a mask (no-mask condition). As shown in Figure 1, both versions included a female instructor standing on the left and slides that illustrated the process of lightning formation on the right. Both versions were recorded by a female instructor with a neutral expression and had the same narrated script, lasting approximately 190 s and containing 755 words. The instructor stood still and looked straight into the camera, which means there were no gaze shifts or gestures. Each version was recorded five times, in which the instructor was required to use the same tone and behave the same each time. Three research assistants chose one recording for each condition to make sure the tone and posture were most similar. Both instructional lessons were created with a screen size of 1280 × 720 pixels. The only difference between the two groups was that the instructor wore a mask in the mask group but did not in the no-mask group.

The prequestionnaire included two parts. The first part collected participants' demographic information (e.g., gender and age). The second part contained a self-rated prior knowledge questionnaire, including nine items ($\omega = 0.92$, e.g., "I know how lightning is formed."), in which participants respond to each item from *1 = very little* to *5 = very much*. The total score was transformed into a percentage. In the current study, McDonald's Omega (ω), one of the best alternatives for estimating reliability, was used to measure reliability of all questionnaires and tests (Hayes & Coutts, 2020; Trizano-Hermosilla & Alvarado, 2016).

The postquestionnaire measured participants' extraneous cognitive load. Extraneous cognitive load ($\omega = 0.91$) was measured by a 4-item 5-point Likert-type subscale (e.g., "I do not like the presentation or organization of the learning material.") from *1 = completely disagree* to *5 = completely agree* (Chang et al., 2017). Scores were summed for analysis.

The post-tests contained the retention test and transfer test ($\omega = 0.51$). Except for the instructions for the retention test, both tests were the same as Moreno and Mayer (1999). The retention test asked participants to describe the steps of lightning formation with as much detail as possible (e.g., "Review what you just learned and write as much detail as possible about how lightning is formed."). Regardless of specific wording, participants received one point for each correct information unit, with a maximum of 19 points in the retention test. The transfer test contained four questions (e.g., "What could you do to decrease the intensity of lightning?"). Each question had a maximum score of two points and the total test had a maximum score of eight points. Each correct expression in an answer received one point. Two raters independently rated the scores and inter-rater reliabilities between two raters were 0.99 and 0.98 ($ps < .001$), respectively, for the retention test and transfer test. The mean scores were used in the analysis.

2.1.3 | Apparatus

Participants' eye movements during learning were recorded using an SMI RED 250 Desktop eye-tracker (SensoMotoric Instruments, Teltow, Germany) with a sampling rate of 250 Hz and a spatial resolution of less than 0.1°. The videos were presented on the monitor with 1680 × 1050 pixels resolution. Participants were seated at around 65 cm from the monitor with their heads fixed on the chinrest and took a headphone to listen to the audio. The fixation filtering threshold was set at 100 ms. We used a 9-point calibration procedure, which is commonly used calibration method (Blignaut et al., 2014). All participants' deviation was below 0.5°. When analyzing eye movement data, we excluded five participants whose eye tracking ratio is low (<75%). The mean tracking ratio is 93.06% (SD = 4.84) in the mask condition and 92.97% (SD = 5.18) in the no-mask condition, which is acceptable.

2.1.4 | Gaze measures

In the study, we assessed participants' gaze behavior during the learning phase with eye movements. We defined three Areas of Interest

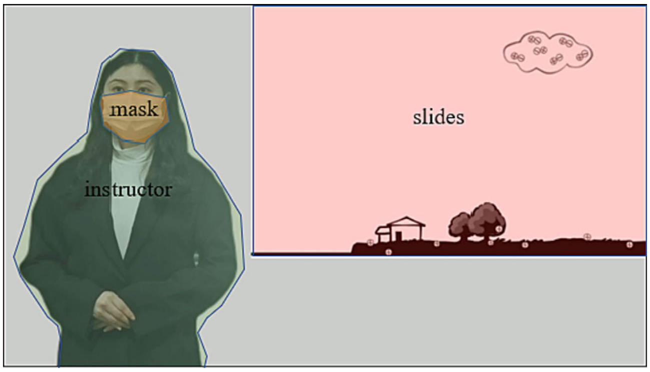

FIGURE 2 There areas of interests in Experiment 1.

(AOIs): slides, instructor, and mask (i.e., bottom half of face region) as shown in Figure 2. The AOI of slides covers 43% of whole screen, the AOI of instructor covers 19%, and the AOI of mask covers 0.9%. Absolute dwell time, relative dwell time, absolute fixation time, and relative fixation time were chosen to indicate how much participants focused their visual attention on each of the three AOIs. Dwell time and fixation time are widely used in the studies of pedagogical agents and the split attention effect (e.g., Pi et al., 2019; Pouw et al., 2019; Wang et al., 2018; Zhang et al., 2020).

2.1.5 | Procedure

Participants were tested individually in a psychology lab room. First participants finished the prequestionnaire. Then they were randomly assigned into two groups. All participants sat in front of the screen with a chinrest. After calibration within a 9-point calibration procedure, all participants watched a video. Participants in the mask group watched the video of the instructor wearing the mask, and those in the no-mask group watched the other video. After that, all participants completed the postquestionnaire (i.e., extraneous cognitive load) and then took the post-test at their own rate. The whole procedure lasted approximately 30 min.

2.2 | Results

Table 1 shows the means and standard deviations for each group on measures of learning outcome, measures of visual attention, and subjective ratings of cognitive load.

2.2.1 | Learning outcomes: Do the groups differ on retention and transfer test score?

Hypothesis 1 is that the no-mask group will perform better on the retention and transfer tests than the mask group. Two *t*-tests were conducted to compare learning outcomes (i.e., retention scores and transfer scores) of the two groups. For the retention test, there was no significant difference between the mask group and no-mask group, $t(59) = 1.35$, $p = .183$, $d = 0.03$. For the transfer test, there also was no significant main effect of the instructor's wearing a mask, $t(59) = 1.12$, $p = .268$, $d = 0.02$. Hypothesis 1 was not supported.

2.2.2 | Visual attention: Do the groups differ on total fixation time on the instructor, mask, and slides?

Hypothesis 2 is that the mask group will spend more time looking at the slide and less time looking at the instructor and mask than the no-mask group. Because of non-normal distributions, gaze measures on the instructor and fixation time on the mask region were log-transformed. Then, *t*-tests on gaze measures were performed to assess whether visual attention was differently allocated between the two groups. For eye gaze on slides, there no significant difference between the two groups on absolute dwell time, $t(54) = 1.89$, $p = .064$, $d = 0.51$, relative dwell time, $t(54) = 1.87$, $p = .067$, $d = 0.50$, fixation time, $t(54) = 1.41$, $p = .116$, $d = 0.38$, and relative fixation time, $t(54) = 1.39$, $p = .170$, $d = 0.37$. For all gaze measures on the instructor, there was a significant main effect of wearing a mask (dwell time: $t(54) = 3.54$, $p = .001$, $d = 0.88$; relative dwell time: $t(54) = 3.55$, $p = .001$, $d = 0.88$; fixation time: $t(54) = 3.55$, $p = .001$, $d = 0.89$; relative fixation time: $t(54) = 3.54$, $p = .001$, $d = 0.89$), indicating more visual attention on the instructor in the no-mask group. Finally, there also was a significant main effect of wearing a mask, showing more attention on the bottom half of the face when the instructor did not wear a mask (dwell time: $t(54) = 3.57$, $p = .001$, $d = 1.16$; relative dwell time: $t(54) = 4.68$, $p < .001$, $d = 1.16$; fixation time: $t(54) = 3.57$, $p = .001$, $d = 1.15$; relative fixation time: $t(54) = 4.65$, $p < .001$, $d = 1.15$). These findings partially support hypothesis 2, indicating that instructor mask wearing affects visual attention in which learners spend less time looking at instructors who wear masks.

2.2.3 | Subjective perception: Do the groups differ on ratings of extraneous load?

To explore whether participants' extraneous load ratings differ in the two groups, a *t*-test was carried out. Results showed that there were no significant differences on extraneous cognitive load, $t(59) = 0.27$, $p = .789$, $d < 0.01$. Hypothesis 3 was not supported.

2.3 | Discussion

Results supported the split attention hypothesis, indicating that the instructor displaying their full face would draw learners' visual attention. As stated before, human faces automatically draw listeners' attention (Gullberg & Holmqvist, 2006). In the multimedia learning environments, such an automatic process caused split attention between the instructor and slides, reflecting more fixation time on the

Variables	Mask	No-mask	t	p
Prior knowledge (%)	32.87 (17.54)	33.33 (19.25)	0.10	.922
Visual attention				
Dwell time on whole screen (s)	162.24 (15.95)	161.60 (18.54)	0.14	.889
Relative dwell time on whole screen (%)	84.97 (8.35)	84.72 (9.71)	0.10	.917
Fixation time on whole screen (s)	137.82 (25.86)	140.42 (24.84)	0.38	.703
Relative fixation time on whole screen (%)	72.14 (13.55)	73.63 (13.02)	0.41	.683
Dwell time on slides (s)	120.81 (18.76)	108.56 (28.41)	1.89	.064
Relative dwell time on slides (%)	63.27 (9.84)	56.92 (14.91)	1.87	.067
Fixation time on slides (s)	105.67 (22.90)	95.86 (28.79)	1.41	.116
Relative fixation time on slides (%)	55.34 (12.00)	50.25 (15.09)	1.39	.170
Dwell time on instructor (s)	19.09 (12.36)	36.69 (25.36)	3.54	.001
Relative dwell time on instructor (%)	10.00 (6.46)	19.23 (13.29)	3.55	.001
Fixation time on instructor (s)	17.49 (11.73)	34.79 (24.76)	3.55	.001
Relative fixation time on instructor (%)	9.16 (6.14)	18.24 (12.99)	3.54	.001
Dwell time on mask (s)	7.53 (8.17)	27.69 (23.22)	3.57	.001
Relative dwell time on mask (%)	3.94 (4.27)	14.51 (12.18)	4.68	<.001
Fixation time on mask (s)	7.30 (7.91)	26.99 (22.85)	3.57	.001
Relative fixation time on mask (%)	3.82 (4.15)	14.14 (11.99)	4.65	<.001
Learning outcomes				
Retention (0–19)	6.63 (2.98)	5.71 (2.36)	1.37	.183
Transfer (0–8)	3.10 (1.52)	3.52 (1.39)	1.12	.268
Subjective rating				
Extraneous load (4–20)	12.40 (4.37)	12.10 (4.44)	0.27	.789

TABLE 1 Means and standard deviations of measures in Experiment 1.

bottom of the face when the instructor did not wear a mask (Ayres & Sweller, 2021). When wearing a mask, dynamic face movements were covered, resulting in less attention to the instructor's face. However, hypotheses 1 and 3 were not supported. Students obtained similar levels of learning outcome performance and perceived extraneous load.

The reason might be that two mechanisms were at work. On one hand, masks could help learning by causing learners to spend less time looking at the instructor's face. On the other hand, masks could hurt learning by obscuring useful cognitive and social cues. These counteracting mechanisms could cancel each other out.

However, we cannot conclude the two mechanisms worked together or actually had any effects. Therefore, in the next experiment, we tried to separate the two mechanisms by using a 2 (mask, no-mask) × 2 (slides, no-slides) between-subject design. In the slides condition, effects of mask and split attention both exist, while in the no-slides condition, there was only the effect of wearing masks. If both mechanisms worked, there would be an interaction effect in which the split-attention effect disappears in the no-slides condition. If split attention is not damaging, negative impacts of face masks would be equivalent, indicated by no interaction effect.

To find out whether wearing a mask and split attention have an impact on learning performance, we used 2 × 2 between-subject design to investigate whether an interaction effect exists in the Experiment 2. To generalize results, in Experiment 2, we used different instructional material involving the water cycle. Additionally, we assessed participants' ratings of affective evaluation and social presence.

3 | EXPERIMENT 2

3.1 | Method

3.1.1 | Participants and design

A prior analysis of G*Power 3.1 suggested a minimum of 52 participants when using a large effect size $f = 0.40$ with power set to 0.80 and alpha set to .05 (Faul et al., 2007). One hundred and one participants were recruited from Central China Normal University. All participants had normal or corrected-to-normal visions. Their mean age was 20.38 years ($SD = 1.71$) and there were 85 women and 16 men. The study was also approved by the ethics committee.

The experiment used a 2 (mask vs. no-mask) × 2 (slides vs. no-slides) between-subject design. Participants were randomly assigned one of to four groups: 25 served in the mask/slide group (instructor wearing a mask and having slides), 26 in the no-mask/slide group (instructor not wearing a mask but having slides), 25 in the mask/no-slide group (instructor wearing a mask but not having slides), and 25 in the no-mask/no-slide group (instructor not wearing a mask and not having slides). There were no significant differences among the groups in prior knowledge, $F(3, 97) = 0.98$, $p > .05$; mean age, $F(3, 97) = 1.31$, $p > .05$; and proportion of women and men ($\chi^2(3) = 1.29$, $p > .05$).

3.1.2 | Materials

The instructional materials were adapted from a course on *icourse* platform (https://www.icourse163.org), a platform of Massive Open Online Courses (MOOCs), and explained the water cycle including the process, rules, principles, and factors involved in the water cycle. There were four versions of the video lessons (as exemplified in Figure 3). Video lessons in the mask/slides and no-mask/slides groups included a masked or unmasked female instructor standing on the left and slides about the water cycle on the right. In the mask/no-slides and no-mask/no-slides groups, there was only the masked or unmasked female instructor, and slides were not shown. In all four versions, the instructor kept a neutral facial expression and had the same script, lasting for around 149 s and having 691 words. All instructional lessons were created with a screen size of 1280 × 720 pixels. The recording process was the same as in Experiment 1.

The prequestionnaire consisted of two parts. In the first part, participants reported their demographic information (e.g., gender and age). The second part contained 10 multiple-choice questions to determine participants' prior knowledge ($\omega = 0.56$). Each question had four choices and only one choice was correct. Participants received one point for each correct answer, with a total possible score of 10.

The postquestionnaire measured participants' intrinsic and extraneous cognitive load, affective evaluation, and social presence. Affective evaluation was measured by one item in which participants rated how much they liked the presented video ("How much do you like the video you just watch?") from *1 = do not like* to *9 = like*. Intrinsic cognitive load ($\omega = 0.94$) was measured by a 4-item subscale (e.g., "The learning content is not easy to be comprehended.") with ratings on a 5-point Likert-style scale ranging from *1 = completely disagree* to *5 = completely agree* (Chang et al., 2017). The measure of extraneous load was the same as in Experiment 1. Social presence was measured by a 5-item scale ($\omega = 0.89$, e.g., "It feels like I deal with a 'real' person and not with an imaginary teacher in the video.") with ratings on a 10-point Likert-style scale ranging from *1 = completely disagree* to *10 = completely agree* (Gorter, 2021). The sum of each measure was used in the data analysis. Participants answered the postquestionnaire in the order of affective evaluation, intrinsic cognitive load, extraneous cognitive load, and social presence.

The post-tests included a retention test ($\omega = 0.59$) and a transfer test ($\omega = 0.69$). The retention test had two questions that required participants to describe what they learned (e.g., "Based on what you just learned, please write the process of the water cycle in as much detail as possible."). Regardless of wording, participants received one point for each correct information unit, with a maximum of 15 points

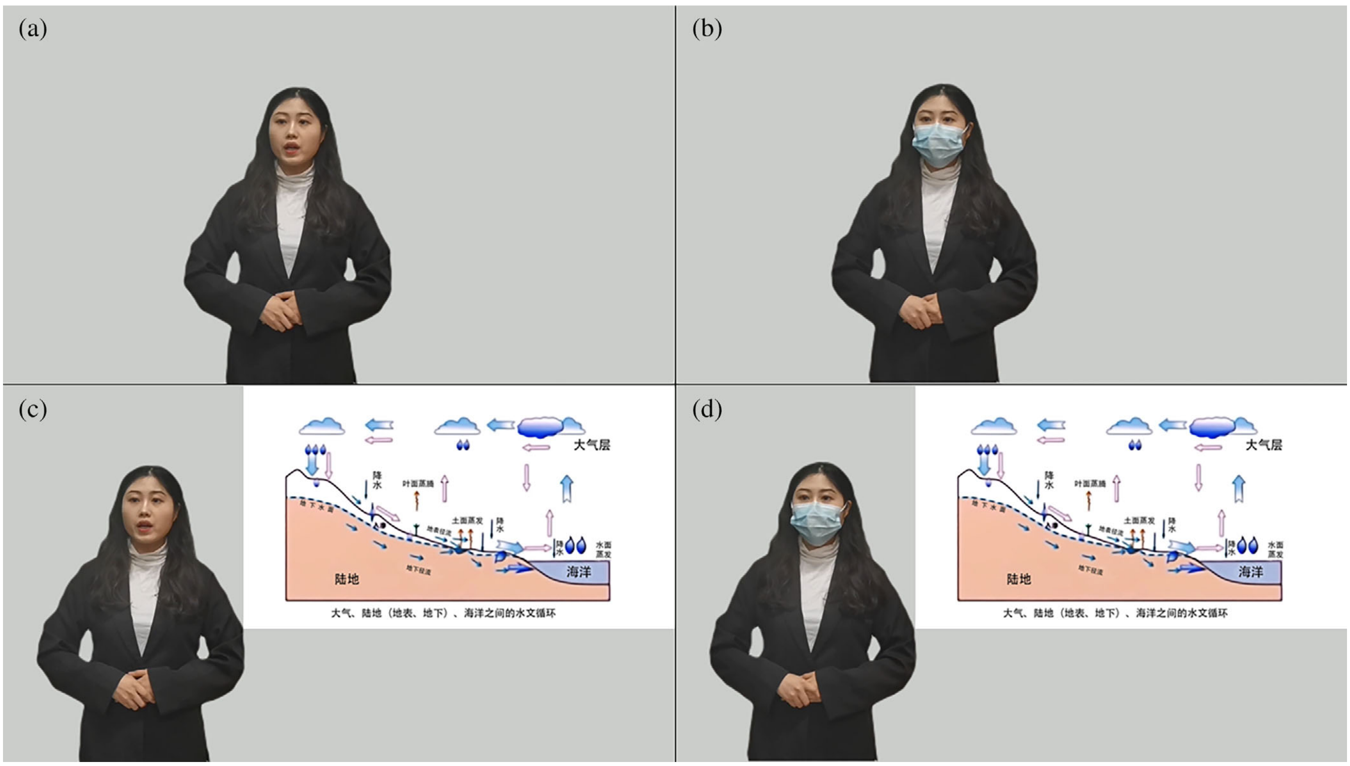

FIGURE 3 Example from the video lessons on the water cycle. Panel a: An example of the no-mask/no-slide group; Panel b: An example of the mask/no-slide group; Panel c: An example of the no-mask/side group; and Panel d: An example of the mask/slide group.

TABLE 2 Means and standard deviations of measures in Experiment 2.

Variables	Slide: Mask	Slide: No-mask	No-slide: Mask	No-slide: No-mask
Prior knowledge (0–10)	5.28 (1.37)	4.96 (1.55)	4.60 (1.63)	5.12 (1.39)
Learning processes				
Social presence (10–50)	24.36 (7.54)	28.62 (9.09)	23.44 (9.73)	26.68 (6.82)
Affective evaluation (1–9)	6.04 (1.43)	5.88 (1.37)	5.28 (1.65)	5.36 (1.66)
Learning outcomes				
Retention (0–15)	9.42 (2.37)	8.10 (3.08)	6.24 (2.34)	7.46 (2.44)
Transfer (0–21)	8.88 (3.01)	8.23 (3.51)	6.92 (2.79)	7.94 (3.09)
Subjective ratings				
Intrinsic load (5–20)	11.00 (4.14)	12.65 (4.13)	16.12 (4.26)	13.92 (5.96)
Extraneous load (5–20)	12.12 (4.56)	11.88 (4.19)	16.64 (5.12)	16.04 (5.36)

(6 points for the first question, and 9 points for the second question). The transfer test contained three open-ended questions (e.g., "The Sanjiang Plain has changed from the 'Great Northern Wilderness' to the 'Great Northern Okura'. What are the water cycle links most affected by human beings and how do they affect?"). Each acceptable expression was worth one point. A maximum of 21 points could be achieved (5 points for the first question, 9 points for the second, and 7 points for the third). Two raters independently rated the answers and the mean scores were used in the analysis. The inter-rater reliabilities between the two raters were 0.97 for the retention test and 0.98 for the transfer test (*ps* $< .001$).

3.1.3 | Procedure

Participants were tested individually in a psychology lab room. Participants first finished the prequestionnaire, and then were randomly assigned into one of the four groups. Participants in the four conditions watched the corresponding video. After that, they completed the postquestionnaire and then the posttest at their own rate. The whole procedure lasted for approximately 30 min.

3.2 | Results

Table 2 shows the means and standard deviations of the four groups on measures of learning process (i.e., social presence rating and affective evaluation), learning outcome (i.e., retention and transfer test score) and subjective ratings of cognitive load. All measures meet the assumptions for conducting ANOVA.

3.2.1 | Learning outcomes: Do the groups differ on retention and transfer test scores?

Hypothesis 1 is that wearing a mask hurts posttest performance. Hypothesis 6 states that adding slides improves learning performance. Hypothesis 7 is that the negative impacts of mask wearing would be less in the slide conditions than in the no slides conditions. Two 2 (mask, no-mask) × 2 (slides, no-slides) ANOVAs were conducted to detect two main effects and the interaction effect. For the retention test, there was no main effect of wearing a mask, $F(1, 97) = 0.01$, $p = .982$, $\eta_p^2 < 0.01$. Hypothesis 1 was not supported. There was a significant main effect of slide displaying, $F(1, 97) = 14.12$, $p < .001$, $\eta_p^2 = 0.13$, that is, participants performed better when watching videos with slides than with no slides, supporting hypothesis 6. This is consistent with the multimedia principle that people learn better from words and graphics than from words alone (Mayer, 2021). In addition, there was a significant interaction effect, $F(1, 97) = 4.28$, $p = .041$, $\eta_p^2 = 0.04$. Results showed that there was a significant effect of displaying slides when the instructor wore a mask, that is, learners' performance in the slide condition was better than in the no slide condition, $F(1, 97) = 16.80$, $p < .001$; when the instructor did not wear a mask, there was no significant difference between the slide and no slide conditions, $F(1, 97) = 1.44$, $p = .233$. Results found an interaction effect, though not as expected.

For the transfer test, in contrast to our hypotheses, there was no main effect of slide displaying, $F(1, 97) = 3.53$, $p = .063$, or of wearing a mask, $F(1, 97) = 0.05$, $p = .828$, and there was no significant interaction effect, $F(1, 97) = 1.72$, $p = .192$.

3.2.2 | Learning processes: Do the groups differ on ratings of social presence and affective evaluation?

Hypothesis 4 is that learners will feel less social presence with the instructor when she wears a mask. For social presence, there was a significant main effect of wearing a mask, $F(1, 97) = 5.05$, $p = .027$, $\eta_p^2 = 0.05$, indicating that learners perceived a higher level of social presence when instructors presented their whole face than when instructors wore masks. There was no significant main effect of displaying slides, $F(1, 97) = 0.73$, $p = .394$, and a no significant interaction, $F(1, 97) = 0.09$, $p = .762$. This supports hypothesis 4.

Hypothesis 5 is that learners will feel less affective evaluation when the instructor wears a mask. For affective evaluation, there was a significant main effect of slide displaying, $F(1, 97) = 4.46$, $p = .037$, $\eta_p^2 = 0.04$, showing that displaying slides increased participants' affective evaluation for the videos. There was neither significant main effect of wearing a mask, $F(1, 97) = 0.02$, $p = .902$, nor interaction effect, $F(1, 97) = 0.15$, $p = .700$. This does not support hypothesis 5.

These findings provide partial support for the idea that mask wearing affects learning processes, at least with respect to social presence.

3.2.3 | Subjective perception

To examine group differences on subjective ratings of cognitive load, two ANOVAs were performed. For intrinsic cognitive load, there was a significant main effect of slide displaying, $F(1, 97) = 11.74$, $p < .001$, $\eta_p^2 = 0.11$, showing that displaying slides decreased intrinsic cognitive load. This supports hypothesis 6. Results also found a significant interaction effect, $F(1, 97) = 4.28$, $p = .041$, $\eta_p^2 = 0.04$: when the instructor wore a mask, learners perceived a higher level of intrinsic cognitive load in the no slide condition than in the slide condition, $F(1, 97) = 14.95$, $p < .001$, but when the instructor did not wear a mask, there was no significant difference between the slide and no slide conditions, $F(1, 97) = 0.93$, $p = 0.337$. Finally, there was no significant main effect of wearing a mask, $F(1, 97) = 0.09$, $p = .770$.

For extraneous cognitive load, results found a significant main effect of slide displaying, $F(1, 97) = 20.42$, $p < .001$, $\eta_p^2 = 0.17$, indicating displaying slides decreased extraneous cognitive load. There was no significant main effect of wearing a mask, $F(1, 97) = 0.19$, $p = .664$, or interaction effect, $F(1, 97) = 0.04$, $p = .850$. Hypothesis 3 was not supported.

3.3 | Discussion

Consistent with Experiment 1, hypotheses 1 and 3 were not supported. Results found no difference on learning tests or extraneous cognitive load between mask and no-mask conditions. Hypothesis 4 was supported, indicating an unmasked instructor raised learner's sense of social presence more than a masked instructor, which was in line with social agency theory (Mayer & DaPra, 2012; Wang et al., 2022). However, hypothesis 5 was not supported, since there was no effect of wearing a mask on affective evaluation. Supporting hypothesis 6, adding slides improved learning outcomes and decreased intrinsic cognitive load. The beneficial effect of slides was consistent with previous studies (e.g., Pi & Hong, 2016). In addition, results found a significant interaction effect between wearing a mask and displaying slides, although not as expected. Adding slides improved learning outcomes only with masked instructor, but not unmasked instructor. Results on intrinsic cognitive load showed similar patterns. When the instructor wore a mask, slides effectively decreased intrinsic cognitive load, and thus, improved learning. In the no-mask conditions, even without slides, the instructor's narration and facial cues provided sufficient information to learn effectively; therefore, adding slides did not significantly decrease intrinsic cognitive load and facilitate learning.

4 | GENERAL DISCUSSION

4.1 | Empirical contributions

The aim of the study was to investigate the impact of the instructor's wearing a mask on students' learning in video-based instruction. In Experiment 1, we hypothesized face masks would impair learning outcomes (hypothesis 1), cause less attention to instructor's face (hypothesis 2), and reduce extraneous load (hypothesis 3). Contrary to hypothesis 1, we found that learning performance was not influenced by the instructor wearing a mask. Hypothesis 2 was supported. Eye movement results showed that students paid more attention to the bottom half of the face region when the instructor showed their whole face than when the instructor wore a mask. In this case, learners allocated their attention on the instructor and slides separately, and more on the instructor's face when there was no mask; however, this was not detrimental to learning. This finding is consistent with previous studies, which indicated the instructor's face would draw learners' attention but not necessarily affect learning outcomes (e.g., Kizilcec et al., 2014; van Gog et al., 2014). However, the on-screen presence of the instructor's face can act as a distractor, so covering faces may help learners to focus on learning content. In contrast, as van Gog et al. (2014) showed, the instructor's face can help learners switch their attention to learning content efficiently, because faces can convey cognitive cues (e.g., lipreading). Contrary to hypothesis 3, face masks did not have an impact on extraneous cognitive load.

In Experiment 2, we hypothesized face masks would impair learning outcomes (hypothesis 1) and reduce extraneous load (hypothesis 3), social presence (hypothesis 4), and affective evaluation (hypothesis 5), and adding slides would improve learning and decrease intrinsic load (hypothesis 6). We also hypothesized that the negative impacts of masking would be less in the slide conditions than in the no slides conditions (hypothesis 7). Consistent with Experiment 1, we found there were no differences on learning outcomes and extraneous load between mask and no-mask conditions, which did not support hypotheses 1 and 3. This result on extraneous cognitive load is consistent with some previous evidence shows no strong impact of onscreen pedagogical agents on extraneous cognitive load (Schroeder & Adesope, 2014). Schroeder and Adesope (2014) argued that results concerning the effects of agents on extraneous cognitive load were mixed, and such mixed results might due to learners, agents' features, or learning materials. In addition, Davis's (2018) meta-analysis found only a minimal effect of gesturing pedagogical agents on cognitive load when compared with control conditions ($g = 0.13$). We found there was a difference on social presence, which supported hypothesis 4, indicating that students felt a higher social presence with the instructor when there was no mask. However, the effect size of social presence is relatively small, so this result should be interpreted cautiously. This result is also consistent with a study of Sondermann and Merkt (2023). Hypothesis 5 was not supported since there was no difference on affective evaluation. The reason might be there was only one item to measure affective evaluation. Results found adding slides improved learning and decreased intrinsic load when the instructor wore a mask, while learning outcomes and intrinsic load remained the same in two no-mask conditions, which partially supported hypothesis 6.

The interaction effect in Experiment 2 indicated negative impacts of face masks, which partially supported hypothesis 7. In the mask conditions, students perceived a higher level of intrinsic cognitive load and performed worse on the retention test when slides were presented than not, while in the no-mask conditions, the effect of adding slides disappeared. Under the mask/no-slides condition, the instructor's bottom half of the face region was covered, and students could only receive very few cognitive cues and social cues, and meanwhile, there were no slides for students, increasing the difficulty of the materials. In contrast, in the mask/slides condition, students shifted their attention to relevant information presented on slides (63% dwell time on slide and only 10% dwell time on the instructor), and intrinsic cognitive load decreased. Overall, in the two no-mask conditions, facial cognitive and social cues provided sufficient information for comprehension; therefore, adding slides did not significantly improve learning. There was no significant interaction effect on transfer test. It might due to the overall low mean score (8 out of 21). However, even though the mean score is low, it cannot be seen as a floor effect.

4.2 | Theoretical implications

To the best of our knowledge, our study was the first to investigate the impact of wearing a mask on learning in educational settings. According to the split attention hypothesis, the instructor's face would draw students' attention, causing split attention (e.g., Kizilcec et al., 2014; Wang & Antonenko, 2017). Results supported this hypothesis, showing more fixation on the face when the instructor presented the whole face. However, it should be noted that, in this case, split attention was not detrimental since there was no direct evidence that it impaired learning.

From the perspective of speech perception, occluding the lower half of face impairs speech perception (e.g., Jordan & Thomas, 2011). Meanwhile, according to social agency theory, facial movements provided social cues which could prime a social stance in learners and make them engage in the organization and integration processes (Fiorella & Mayer, 2022; Mayer, 2021; Mayer & DaPra, 2012). From the two perspectives, students would benefit more from an instructor who was not masked. However, the results were not consistent with these theories in that mask wearing did not affect learning outcomes. Apparently, there was enough information for learners even when they viewed videos with the masked instructors and slides.

4.3 | Practical implications

The most important practical implication is that an instructor wearing a mask does not influence students' video learning. When recording offline lectures or teaching through online platforms, instructors are able to wear masks to isolate the virus without harming students' learning outcomes. In addition, when recording from locations where people have to wear masks, such as chemical labs or hospitals, it is not a problem to wear a mask because it will not impair learning. If instructors want to wear masks during lectures, they should use slides to support learning. We recommend caution in generalizing the results, since we only explored the effect of wearing a mask in video-based instruction, but not in other educational settings such as real classrooms.

4.4 | Limitations and future directions

This study has some limitations. First, we only used one item to assess participants' affective evaluation, which might be the reason why we did not find a significance difference of wearing a mask. Future research could use validated assessments with multiple items to examine how students' affective evaluation is influenced by instructors wearing masks.

Second, we did not distinguish students with high and low prior knowledge. Some research has provided empirical evidence that learners with low prior knowledge would benefit from instructor presence, but learners with high prior knowledge performed better without instructor presence (e.g., Johnson et al., 2015). Thus, instructors' wearing masks may be also moderated by prior knowledge. Future research could examine whether prior knowledge moderates the effect of mask wearing.

Third, the studies were carried out in multimedia environments, so results cannot be simply generalized to all educational settings. In a multimedia environment, many interfering factors in real classroom environments, like noise and the distance between students and instructors, were absent. These factors might mediate the effect of wearing masks. For example, lip movements were demonstrated to be more beneficial in noisy environments since people could hardly hear speakers (e.g., Ma et al., 2009; Vatikiotis-Bateson et al., 1998); therefore, masking the bottom half of the instructor's face might damage learning in noisy classrooms. Future research could examine whether type of learning environment mediates the effect of wearing a mask.

Fourth, in this study, videos in the mask and no-mask conditions were recorded separately. As a consequence, there might be slight differences among those recordings which influenced the impact of wearing a mask. Future research could put a mask on the instructor in a video by using AI technology to ensure similarity among recordings.

Fifth, the instructor in the study just stood alone, and no gaze direction or gestures were involved. Since gaze or gestures can act as signals to guide learners' attention and improve learning (e.g., Li et al., 2019; Pi et al., 2019), future research could test whether adding gaze or gestures will mediate the effect of wearing a mask.

5 | CONCLUSIONS

The aim of this study was to investigate the impact of the instructor's wearing a mask on students' learning in educational settings. Across two experiments, we found that whether or not the instructor wore a standard mask did not influence learning outcomes, but influenced visual attention. Adding a mask lowered student ratings of social presence with the instructor. Additionally, there was a significant interaction between wearing a mask and displaying slides in which when instructor wore a mask, adding slides led to better retention test performance, but when there was no mask, students did not benefit from slides. In conclusion, results supported the idea that there was more fixation on the instructor's face when there was no mask.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (Grant number: 62277025).

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interest.

DATA AVAILABILITY STATEMENT

The data and study materials for this study are available from the corresponding author upon reasonable request.

ORCID

Xiaoxue Leng https://orcid.org/0000-0001-9808-5118
Fuxing Wang https://orcid.org/0000-0003-3373-7095
Richard E. Mayer https://orcid.org/0000-0003-4055-6938

REFERENCES

Ayres, P., & Sweller, J. (2021). The split-attention principle in multimedia learning. In R. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning (Cambridge handbooks in psychology)* (pp. 199–211). Cambridge University Press. https://doi.org/10.1017/9781108894333.020

Blignaut, P., Holmqvist, K., Nyström, M., & Dewhurst, R. (2014). Improving the accuracy of video-based eye tracking in real time through post-calibration regression. In M. Horsley, M. Eliot, B. Knight, & R. Reilly (Eds.), *Current trends in eye tracking research*. Springer, Cham. https://doi.org/10.1007/978-3-319-02868-2_5

Buchan, J. N., Paré, M., & Munhall, K. G. (2007). Spatial statistics of gaze fixations during dynamic face processing. *Social Neuroscience*, *2*(1), 1–13. https://doi.org/10.1080/17470910601043644

Buchan, J. N., Paré, M., & Munhall, K. G. (2008). The effect of varying talker identity and listening conditions on gaze behavior during audio-visual speech perception. *Brain Research*, *1242*, 162–171. https://doi.org/10.1016/j.brainres.2008.06.083

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C., McGuire, P. K., Woodruff, P. W., Iversen, S. D., & David, A. S. (1997). Activation of auditory cortex during silent lipreading. *Science*, *276*(5312), 593–596. https://doi.org/10.1126/science.276.5312.593

Chang, C. C., Liang, C., Chou, P. N., & Lin, G. Y. (2017). Is game-based learning better in flow experience and various types of cognitive load than non-game-based learning? Perspective from multimedia and media richness. *Computers in Human Behavior*, *71*, 218–227. https://doi.org/10.1016/j.chb.2017.01.031

Colliot, T., & Jamet, É. (2018). Understanding the effects of a teacher video on learning from a multimedia document: An eye-tracking study. *Educational Technology Research and Development*, *66*, 1415–1433. https://doi.org/10.1007/s11423-018-9594-x

Davis, C., & Kim, J. (2006). Audio-visual speech perception off the top of the head. *Cognition*, *100*(3), B21–B31. https://doi.org/10.1016/j.cognition.2005.09.002

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

Erber, N. P. (1975). Auditory-visual perception of speech. *The Journal of Speech and Hearing Disorders*, *40*(4), 481–492. https://doi.org/10.1044/jshd.4004.481

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Fiorella, L. (2022). Multimedia learning with instructional video. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 487–497). Cambridge University Press. https://doi.org/10.1017/9781108894333.050

Fiorella, L., & Mayer, R. (2022). Principles based on social cues in multimedia learning: Personalization, voice, image, and embodiment principles. In R. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 277–285). Cambridge University Press. https://doi.org/10.1017/9781108894333.029

Giordano, B. L., Ince, R. A. A., Gross, J., Schyns, P. G., Panzeri, S., & Kayser, C. (2017). Contributions of local speech encoding and functional connectivity to audio-visual speech perception. *eLife*, *6*, e24763. https://doi.org/10.7554/eLife.24763

Gorter, M. (2021). *An investigation of the effect of a pedagogical agent on achievement, motivation, affect and learning for primary school students* (Master thesis). Open Universiteit, Netherlands.

Gullberg, M., & Holmqvist, K. (2006). What speakers do and what addressees look at: Visual attention to gestures in human interaction live and on video. *Pragmatics & Cognition*, *14*, 53–82. https://doi.org/10.1075/pc.14.1.05gul

Haider, C. L., Suess, N., Hauswald, A., Park, H., & Weisz, N. (2022). Masking of the mouth area impairs reconstruction of acoustic speech features and higher-level segmentational features in the presence of a distractor speaker. *NeuroImage*, *252*, 119044. https://doi.org/10.1016/j.neuroimage.2022.119044

Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. But…. *Communication Methods and Measures*, *14*(1), 1–24. https://doi.org/10.1080/19312458.2020.1718629

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

Hyönä, J. (2010). The use of eye movements in the study of multimedia learning. *Learning and Instruction*, *20*(2), 172–176. https://doi.org/10.1016/j.learninstruc.2009.02.013

IJsseldijk, F. J. (1992). Speechreading performance under different conditions of video image, repetition, and speech rate. *Journal of Speech and Hearing Research*, *35*(2), 466–471. https://doi.org/10.1044/jshr.3502.466

Johnson, A. M., Ozogul, G., & Reisslein, M. (2015). Supporting multimedia learning with visual signaling and animated pedagogical agent: Moderating effects of prior knowledge. *Journal of Computer Assisted Learning*, *31*(2), 97–115. https://doi.org/10.1111/jcal.12078

Jordan, T. R., & Thomas, S. M. (2011). When half a face is as good as a whole: Effects of simple substantial occlusion on visual and audiovisual speech perception. *Attention, Perception, and Psychophysics*, *73*, 2270–2285. https://doi.org/10.3758/s13414-011-0152-4

Kizilcec, R. F., Papadopoulos, K., & Sritanyaratana, L. (2014). Showing face in video instruction: Effects on information retention, visual attention, and affect. In *Proceedings of the 32nd annual ACM conference on human factors in computing systems* (pp. 2095–2102). ACM. https://doi.org/10.1145/2556288.2557207

Lansing, C. R., & McConkie, G. W. (1999). Attention to facial regions in segmental and prosodic visual speech perception tasks. *Journal of Speech, Language, and Hearing Research*, *42*(3), 526–539. https://doi.org/10.1044/jslhr.4203.526

Lansing, C. R., & McConkie, G. W. (2003). Word identification and eye fixation locations in visual and visual-plus-auditory presentations of spoken sentences. *Perception & Psychophysics*, *65*, 536–552. https://doi.org/10.3758/BF03194581

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

Lusk, L. G., & Mitchel, A. D. (2016). Differential faze patterns on eyes and mouth during audiovisual speech segmentation. *Frontiers in Psychology*, *7*, 52. https://doi.org/10.3389/fpsyg.2016.00052

Ma, W. J., Zhou, X., Ross, L. A., Foxe, J. J., & Parra, L. C. (2009). Lip-reading aids word recognition most in moderate noise: A Bayesian explanation using high-dimensional feature space. *PLoS One*, *4*(3), e4638. https://doi.org/10.1371/journal.pone.0004638

Marini, M., Ansani, A., Paglieri, F., Caruana, F., & Viola, M. (2021). The impact of facemasks on emotion recognition, trust attribution and re-identification. *Scientific Reports*, *11*, 5577. https://doi.org/10.1038/s41598-021-84806-5

Mayer, R. E. (2010). Unique contributions of eye-tracking research to the study of learning with graphics. *Learning and Instruction*, *20*(2), 167–171. https://doi.org/10.1016/j.learninstruc.2009.02.012

Mayer, R. E. (2021). *Multimedia learning* (3rd ed.). Cambridge University Press.

Mayer, R. E. (2022). The multimedia principle. In R. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 145–157). Cambridge University Press. https://doi.org/10.1017/9781108894333.015

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*, 837–852. https://doi.org/10.1007/s11423-020-09749-6

McCrackin, S. D., Capozzi, F., Mayrand, F., & Ristic, J. (2023). Face masks impair basic emotion recognition: Group effects and individual variability. *Social Psychology*, *54*(1–2), 4–15. https://doi.org/10.1027/1864-9335/a000470

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, *91*(2), 358–368. https://doi.org/10.1037/0022-0663.91.2.358

Park, H., Kayser, C., Thut, G., & Gross, J. (2016). Lip movements entrain the observers' low-frequency brain oscillations to facilitate speech intelligibility. *eLife*, *5*, e14521. https://doi.org/10.7554/eLife.14521

Pi, Z., & Hong, J. (2016). Learning process and learning outcomes of video podcasts including the instructor and PPT slides: A Chinese case. *Innovations in Education Teaching International*, *53*(2), 135–144. https://doi.org/10.1080/14703297.2015.1060133

Pi, Z., Xu, K., Liu, C., & Yang, J. (2019). Instructor presence in video lectures: Eye gaze matters, but not body orientation. *Computers & Education*, *144*, 103713. https://doi.org/10.1016/j.compedu.2019.103713

Pouw, W., Rop, G., de Koning, B., & Paas, F. (2019). The cognitive basis for the split-attention effect. *Journal of Experimental Psychology: General*, *148*(11), 2058–2075. https://doi.org/10.1037/xge0000578

Rennig, J., & Beauchamp, M. S. (2018). Free viewing of talking faces reveals mouth and eye preferring regions of the human superior temporal sulcus. *NeuroImage*, *183*, 25–36. https://doi.org/10.1016/j.neuroimage.2018.08.008

Rinck, M., Primbs, M. A., Verpaalen, I. A. M., & Bijlstra, G. (2022). Face masks impair facial emotion recognition and induce specific emotion confusions. *Cognitive Research: Principles and Implications*, *7*(1), 83. https://doi.org/10.1186/s41235-022-00430-5

Schroeder, N. L., & Adesope, O. O. (2014). A systematic review of pedagogical agents' persona, motivation, and cognitive load implications for learners. *Journal of Research on Technology in Education*, *46*(3), 229–251. https://doi.org/10.1080/15391523.2014.888265

Sommers, M. S., & Phelps, D. (2016). Listening effort in younger and older adults: A comparison of auditory-only and auditory-visual presentations. *Ear and Hearing*, *37*(Suppl 1), 62S–68S. https://doi.org/10.1097/AUD.0000000000000322

Sondermann, C., & Merkt, M. (2023). Like it or learn from it: Effects of talking heads in educational videos. *Computers & Education*, *193*, 104675. https://doi.org/10.1016/j.compedu.2022.104675

Spitzer, M. (2020). Masked education? The benefits and burdens of wearing face masks in schools during the current Corona pandemic. *Trends in Neuroscience and Education*, *20*, 100138. https://doi.org/10.1016/j.tine.2020.100138

Stull, A., Fiorella, L., & Mayer, R. E. (2018). 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

Sweller, J., Ayres, P., & Kalyuga, S. (2011). Intrinsic and extraneous cognitive load. In *Cognitive load theory* (pp. 57–69). Springer. https://doi.org/10.1007/978-1-4419-8126-4_5

Trizano-Hermosilla, I., & Alvarado, J. M. (2016). Best alternatives to Cronbach's alpha reliability in realistic conditions: Congeneric and asymmetrical measurements. *Frontiers in Psychology*, *7*, Article 769. https://doi.org/10.3389/fpsyg.2016.00769

van Gog, T., Verveer, I., & Verveer, L. (2014). Learning from video modeling examples: Effects of seeing the human model's face. *Computer & Education*, *72*, 323–427. https://doi.org/10.1016/j.compedu.2013.12.004

van Wermeskerken, M., Ravensbergen, S., & van Gog, T. (2018). Effects of instructor presence in video modeling examples on attention and learning. *Computers in Human Behavior*, *89*, 430–438. https://doi.org/10.1016/j.chb.2017.11.038

van Wermeskerken, M., & van Gog, T. (2017). Seeing the instructor's face and gaze in demonstration video examples affects attention allocation but not learning. *Computers & Education*, *113*, 98–107. https://doi.org/10.1016/j.compedu.2017.05.013

Vatikiotis-Bateson, E., Eigsti, I. M., Yano, S., & Munhall, K. G. (1998). Eye movement of perceivers during audiovisual speech perception. *Perception & Psychophysics*, *60*(6), 926–940. https://doi.org/10.3758/bf03211929

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. https://doi.org/10.1037/edu0000221

Wang, F., Li, W., & Zhao, T. (2022). Multimedia learning with animated pedagogical agents. In R. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 450–460). Cambridge University Press. https://doi.org/10.1017/9781108894333.047

Wang, J., Antonenko, P., & Dawson, K. (2020). Does visual attention to the instructor in online video affect learning and learner perceptions? An eye-tracking analysis. *Computers & Education*, *146*, 103779. https://doi.org/10.1016/j.compedu.2019.103779

Wang, J., & Antonenko, P. D. (2017). Instructor presence in instructional video: Effects on visual attention, recall, and perceived learning. *Computers in Human Behavior*, *71*, 79–89. https://doi.org/10.1016/j.chb.2017.01.049

Zhang, H., Miller, K. F., Sun, X., & Cortina, K. S. (2020). Wandering eyes: Eye movements during mind wandering in video lectures. *Applied Cognitive Psychology*, *34*(2), 449–464. https://doi.org/10.1002/acp.3632

Zhang, L., & Du, Y. (2022). Lip movements enhance speech representations and effective connectivity in auditory dorsal stream. *NeuroImage*, *257*, 119311. https://doi.org/10.1016/j.neuroimage.2022.119311

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Leng, X., Wang, F., & Mayer, R. E. (2024). Is student learning from a video lecture affected by whether the instructor wears a mask? *Applied Cognitive Psychology*, *38*(1), e4169. https://doi.org/10.1002/acp.4169