Received: 9 April 2018 | Revised: 13 August 2018 | Accepted: 19 August 2018

DOI: 10.1111/jcal.12309

ORIGINAL ARTICLE

WILEY Journal of Computer Assisted Learning

# The instructor's gaze guidance in video lectures improves learning

Hongyan Wang[1,2] | Zhongling Pi[1] | Weiping Hu[1,3]

[1] Key Laboratory of Modern Teaching Technology (Ministry of Education), Center for Teacher Professional Ability Development, Shaanxi Normal University, Xi'an, Shaanxi, China

[2] School of Education Science, Weinan Normal University, Weinan, Shaanxi, China

[3] National Innovation Center for Assessment of Basic Education Quality, Beijing Normal University, Beijing, China

**Correspondence**
Weiping Hu, Shaanxi Normal University, Xi'an 710062, Changannan Road 199, Yanta District, China.
Email: weipinghu@163.com

**Funding information**
the Fundamental Research Funds for the Central Universities, Grant/Award Number: GK201803076; the China Postdoctoral Science Foundation, Grant/Award Number: 2018M631118; the Research Program Funds of the collaborative innovation center of assessment toward Basic Education Quality at Beijing Normal University, Grant/Award Number: 2018-05-002-BZPK01; Teaching Reform Project of Weinan Normal University, Grant/Award Number: JG201706; Educational Technology in Weinan Normal University of First-class Specialty Cultivate Project of Shaanxi Universities, Grant/Award Number: SJ2017-355-425

## Abstract

Instructor behaviour is known to affect learning performance, but it is unclear which specific instructor behaviours can optimize learning. We used eye-tracking technology and questionnaires to test whether the instructor's gaze guidance affected learners' visual attention, social presence, and learning performance, using four video lectures: declarative knowledge with and without the instructor's gaze guidance and procedural knowledge with and without the instructor's gaze guidance. The results showed that the instructor's gaze guidance not only guided learners to allocate more visual attention to corresponding learning content but also increased learners' sense of social presence and learning. Furthermore, the link between the instructor's gaze guidance and better learning was especially strong for participants with a high sense of social connection with the instructor when they learned procedural knowledge. The findings lead to a strong recommendation for educational practitioners: Instructors should provide gaze guidance in video lectures for better learning performance.

KEYWORDS

eye tracking, instructor's gaze guidance, learning, social presence, video lecture

## 1 | INTRODUCTION

Video lectures have increasingly become a part of online teaching and learning (Bos, Groeneveld, van Bruggen, & Brand-Gruwel, 2016). While watching video lectures, learners are required not only to process the information being presented but also to memorize the information that has disappeared from the screen (Mayer & Chandler, 2001). However, compared with traditional classrooms, video lectures do not always have supplementary materials (e.g., the textbook and blackboard writing) that could help learners memorize the course content. As a consequence, learners have reported that they had to be far more focused when watching video lectures than when learning in a traditional classroom in which the instructor and learners are face to face in a shared physical classroom, and there is concern that the effectiveness of video lectures may not be satisfactory in educational practice (O'Bannon, Lubke, Beard, & Britt, 2011; Traphagan, Kucsera, & Kishi, 2010). However, empirical evidence suggests that adding attentional cues, such as symbols, colour text, and gestures, guides learners' visual attention and thus improves their learning performance in video lectures (Johnson, Ozogul, & Reisslein, 2015; F. X. Wang, Li, Mayer, & Liu, 2018).

Attentional cues provide non-content information that can guide learners' visual attention at the right time to the information taught by the instructor (Amadieu, Marine, & Laimay, 2011). According to Mayer's multimedia learning theory, by effectively guiding learners' attention to the essential learning information, attentional cues decrease the competition in working memory resources between relevant and irrelevant information and also decrease unnecessary cognitive load (De & Tabbers, 2013; Mayer, 2005; Ozcelik, Arslan-

Hongyan Wang and Zhongling Pi are the first co-authors.

Ari, & Cagiltay, 2010). As a consequence, adding attentional cues in video lectures should optimize the learner's performance.

Attentional cues are conceptualized as nonhuman cues (provided by the multimedia representation of learning content) and human cues (provided by the instructor's behaviour; Amadieu et al., 2011; Ouwehand, Van Gog, & Paas, 2015). Human gestures appear to improve learning more than nonhuman cues (e.g., De & Tabbers, 2013; Pi, Hong, & Yang, 2017a). This may be because learners need fewer working memory resources to follow the instructor's gestures than to follow nonhuman cues, as gestures could elicit joint attention, even without awareness of the gestures (Pi et al., 2017a).

Human cues in video lectures not only include the instructor's gestures but also include gaze guidance. However, most previous studies have focused only on the instructor's gestures (Alibali et al., 2013; Ouwehand, Van Gog, & Paas, 2015; Pi et al., 2017a; Rueckert, Church, Avila, & Trejo, 2017; Yeo, Ledesma, Nathan, Alibali, & Church, 2017). Gaze guidance in video lectures means that the instructor is not merely looking into the camera but switches between looking at the camera and the content representation domain. These switches might direct the learner's visual attention away from the instructor's image to the information that the instructor is emphasizing at the moment (Gog, Verveer, & Verveer, 2014; Ouwehand, Van, & Paas, 2015).

Furthermore, according to the social presence theory, the instructor's gaze guidance might not only direct learners' attention towards what the instructor is teaching but also provide a sense of social connection and interaction between the instructor and the learner (Gunawardena, 1995; Kizilcec, Papadopoulos, & Sritanyaratana, 2014). Learners' sense of interacting with others in a virtual environment has been called social presence (Gunawardena, 1995). As social presence theory suggested, learners' sense of high social presence stimulates their motivation and learning engagement, and positively predicts learning performance (Lyons, Reysen, & Pierce, 2012; F. X. Wang et al., 2018; Joksimovic, Gasevic, Kovanovic, Riecke, & Hatala, 2015).

Studies on video lectures with the instructor's image have indicated the potential benefits of gaze guidance for learning (Pi & Hong, 2016; J. Wang & Antonenko, 2017). Recently, some studies directly tested the effects of the instructor's gaze on learning (Gog et al., 2014; Ouwehand, Van, & Paas, 2015; Wermeskerken & Gog, 2017). For example, Ouwehand, Van, and Paas (2015) compared learners' attention, learning performance, and mental effort in video lectures under three conditions (i.e., no cues, gaze cues, or gaze and gesture cues). They found that the gaze and gesture cues condition was effective in directing learners' attention from the instructor area to the content areas that the instructor was referring to but produced no positive effects on learning performance or mental effort. There are two other studies that have directly tested the effect of gaze guidance in video lectures (Gog et al., 2014; Wermeskerken & Gog, 2017). In the first study, the researchers compared 43 adult participants' performance after they viewed the "Frog Leap" video lecture in which the teacher's face was either visible or not. The researchers concluded that the teacher's face might provide gaze guidance, helping participants to switch their attention from the teacher's image area to the task area. On the basis on this assumption, Wermeskerken and Gog (2017) studied 69 psychology undergraduates' performance in a video lecture on the composition of glutamine (C5H10N2O3) and construction of the glutamine molecule, under three conditions: the teacher's face was invisible, was visible without gaze guidance, or was visible with gaze guidance. However, in contrast to the "Frog Lea" study, learners who watched the video lecture including the instructor's gaze guidance did not know more about the construction of the molecule than those who watched the video lecture that included the instructor's face but did not include gaze guidance.

The inconsistent results of these two studies are likely due to differences in learning content. For example, in Gog et al.'s (2014) study, the learning material was a puzzle problem-solving task, namely, "Frog Leap." In the video lecture, the instructor's demonstration of each leap step was transient, so paying attention to the right place at the right time was crucial. In this task, the instructor's gaze guidance helped learners to switch their visual attention from the instructor's face to the task area in a timely and efficient manner. However, in Wermeskerken and Gog's (2017) study, the learning content was the composition of glutamine (C5H10N2O3) and construction of the glutamine molecule. The instructor explained some basic characteristics of atoms and bonds and demonstrated how to build the glutamine molecule. This learning content was continuously visible, and in the event that the instructor's face distracted the learners' visual attention, they could catch up with the missed information by analysing the state of the molecule at the moment.

The learning content in Gog et al.'s (2014) study was essentially a lesson in procedural knowledge, and the learning content in Wermeskerken and Gog's (2017) study was essentially a lesson in declarative knowledge. Declarative knowledge and procedural knowledge are different in many aspects. Declarative knowledge could be described as "know-what," such as facts, theories, and names; procedural knowledge could be described as "know-how," such as method, skill, and arithmetic (Anderson, 1982). Mastering declarative knowledge emphasizes remembering and understanding knowledge. Mastering procedural knowledge emphasizes not only remembering and understanding but also mastering the processes of operation. Therefore, the process of learning procedural knowledge is relatively difficult and requires more cognitive resources than the process of learning declarative knowledge (Anderson, 1995).

Furthermore, some studies did show that learning procedural knowledge in video lectures was different from learning declarative knowledge (Hong, Pi, & Yang, 2018). These results imply that research on the factors that influence the effects of video lectures on learning should consider the learning content of what was taught (Winslett, 2014). This raises a new question: Does the instructor's gaze guidance in video lectures have the same positive effects on the learning of procedural knowledge and declarative knowledge? We assumed that learning procedural knowledge would require the learner to obtain key learning content more quickly than learning declarative knowledge, and thus, the instructor's gaze guidance would have stronger positive effects on learning procedural than declarative knowledge.

Our study was designed to examine the visual attention, social presence, and learning performance of learners who watched a video lecture with or without the instructor's gaze guidance, taking into account whether the learning content was declarative or procedural knowledge. Based on the above literature review, our hypotheses were as follows:

H1. For procedural knowledge and declarative knowledge, learners who watch the video lecture with the instructor's gaze guidance will allocate more visual attention to areas of the PowerPoint (PPT) slides and switch less often between areas of the PPT slides and areas of the instructor's image than those who watch the video lecture without the instructor's gaze guidance.

H2. For procedural knowledge and declarative knowledge, learners who watch the video lecture with the instructor's gaze guidance will report a higher level of social presence than those who watch the video lecture without the instructor's gaze guidance.

H3. The instructor's gaze guidance in video lectures will have stronger positive effects on learning procedural knowledge than learning declarative knowledge. Specifically, learners who watch the video lecture with the instructor's gaze guidance will show significantly better learning performance than those who watch the video lecture without the instructor's gaze guidance for procedural knowledge; for declarative knowledge, there will be no difference between the groups.

H4. Learners' level of social presence will moderate the relation between gaze guidance and the learning of procedural knowledge but not declarative knowledge.

# 2 | METHOD

## 2.1 | Participants

Participants were 64 undergraduates (age $M$ = 18.74, $SD$ = 0.77; male = 22). They were recruited from a Chinese university through online advertisements and were required to have used office software for over 1 year at some point. The participants majored in a broad variety of disciplines (e.g., psychology, mathematics, geography, economics, Chinese, English, and physics), and none of them was majoring in educational technology. None of the participants had prior knowledge of the topic presented in the video lectures. All participants were tested for achromatopsia and visual acuity. They all had normal or corrected-to-normal vision. Six participants were excluded due to calibration errors (deviation of calibration exceeding 1°). All participants provided written informed consent. The study protocol was approved by the local ethics committee. The 58 participants were randomly assigned to two conditions: with instructor's gaze guidance ($n$ = 29) and without instructor's gaze guidance ($n$ = 29).

## 2.2 | Materials

The video lectures about declarative knowledge and procedural knowledge were part of an online course for undergraduates majoring in educational technology. The declarative knowledge topic was the red, green, and blue colour model, and the procedural knowledge topic was image composition skill. The declarative knowledge and procedural knowledge video lectures each included two kinds of presentation. (a) In the *with gaze guidance* condition, the instructor stood next to the PPT slides or Photoshop software interface, and she looked at the PPT slides nine times. (b) In the *without gaze guidance* condition, the instructor stood next to the PPT slides or Photoshop software interface, and she always looked at the camera (see Figure 1). The declarative knowledge video lecture (with and without gaze guidance) lasted 6 min and 55 s, and the procedural knowledge video lecture (with and without gaze guidance) lasted 6 min and 54 s. The instructor used the same tone, intonation, description, and explanation of vocabulary in the two videos of each knowledge type. The instructor's


**FIGURE 1** Screenshot of each of video lectures on declarative and procedural knowledge [Colour figure can be viewed at wileyonlinelibrary.com]

hair, clothes, facial expression, and image size on the screen were also the same in all video lectures.

## 2.3 | Apparatus

Eyelink 1000 plus (SR Research Ltd., Canada) was used to collect eye tracker data. A 9-point calibration and validation procedure determined fixation and eye reaction to stimuli presentation before viewing the video lectures. The video lectures were played on a 21-in. monitor. The distance between the CRT monitor and the participants was 60 cm. The resolution of the monitor was 1280 × 768 pixels, and the refresh rate was 75 Hz.

## 2.4 | Measurements

### 2.4.1 | Demographic questionnaire

Participants reported their gender, age, year in school, and major. They also reported how many years they had been using office software.

### 2.4.2 | Prior knowledge test

The prior knowledge test, designed for the current study, was used to assess the participants' prior knowledge of Photoshop software. The test consists of 10 multiple-choice questions with four alternative answers, of which two to four answers are correct. If participants selected all correct answers and rejected all false answers in one question, they were assigned 1 point, with a maximum total score of 10 points. (e.g., Which procedure in Photoshop can be used to copy a layer? A. Select "edit" > copy layer. B. Click the right mouse button > copy. C. Select "image" > copy layer. D. Select "file" > copy layer.) No significant difference in prior knowledge was found between the with gaze guidance and without gaze guidance groups, $t(58) = 0.58$, $p > 0.05$, $MD = 1.43$, suggesting that the two groups had a similar level of prior knowledge.

### 2.4.3 | Visual attention

Two areas of interest (AOI) were created on the video lectures. The top left coordinate point of the video interface was 0, 0. AOI1 was the PPT slides area of interest, (822–2) × (715–2) pixels, which captured the area located in the right part of the video. AOI2 was the instructor area of interest, (1271–831) × (715–283) pixels, which captured the instructor's image located in the bottom left part of the video. To analyse whether students in the *with gaze guidance* condition indeed looked at the PPT slides significantly more than students in the *without gaze guidance* condition, we measured the total fixation duration on the PPT slides area when the instructor provided gaze guidance and saccade counts from the area of the instructor to the PPT slides.

### 2.4.4 | Social presence scale

The social presence scale was developed by Kim and Biocca (1997). We used the Chinese version of the scale (Pi, Hong, & Yang, 2017b). The scale has eight items rated on a Likert-type scale, ranging from 1 (*not at all*) to 5 (*a great deal*;e.g., When I viewed the video lecture, I immersed myself in the virtual learning environment). In the current sample, this scale had high reliability (internal consistency coefficient, Cronbach's $\alpha = 0.868$).

### 2.4.5 | Learning performance test

#### Learning performance test on declarative knowledge

All questions on the declarative knowledge test, designed for the current study, were derived from the content presented in the video lectures. The test included 10 multiple-choice questions (4 points each) and two open-ended questions (10 points each). The maximum total score, across both types of items, was 60 points. The scoring standard for the 10 multiple-choice questions was the same as for the prior knowledge test. For the two open-ended questions, two trained raters who did not know the aim of the study scored the answers on the following dimensions: using the concepts to describe the phenomenon (2 points), stating the rules of colour blending (4 points), and giving the method for solving the colour problem (4 points). The scores on open-ended questions were the average of the two raters' scores. The raters showed high interrater agreement on each of the two questions (respectively, Cronbach's $\alpha$ = 0.75 and 0.87). This test had high discriminant validity, $t(34) = 11.89$, $p < 0.01$, $MD = 26.5$, in that it was able to discriminate between participants who made high versus low gains in learning.

#### Learning performance test on procedural knowledge

On the test of procedural knowledge, designed for the current study, participants were asked to synthesize two pictures to create a new picture using the mask function in Photoshop. The first group of pictures was similar to the example in the video lecture, and the second group of pictures was different from the video lecture example. We recorded the process of participants' synthesizing the two pictures. The study used 15 dimensions to evaluate the videos of the participants' synthesis procedures: selecting picture, opening picture, moving window, moving picture, selecting the picture layer, adding mask for current layer, selecting brush, setting the foreground colour, painting on the mask layer, changing the size of the brush, changing foreground colour to write, setting material layer order, using the other filled tool, modifying the perfect effect, and save format. Each dimension was worth 2 points, meaning 30 possible points for each of the items, with the maximum total score of 60 points for the procedural knowledge learning test. Two independent raters assessed the participants' processing video and showed high interrater agreement on the two tasks (respectively, Cronbach's $\alpha = 0.92$ and 0.87). This test's discriminant validity was high, $t(34) = 14.96$, $p < 0.01$, $MD = 24.2$, as it was able to discriminate between participants who showed high versus low gains in learning.

## 2.5 | Design

The study used a mixed between- and within-subjects design. The presenting mode of the video lectures (with instructor's gaze guidance vs. without gaze guidance) was the between-subjects variable, and the type of knowledge (declarative vs. procedural) was the within-subjects variable.

### 2.6 | Procedure

The study was carried out in an eye-tracking laboratory and took approximately 90 min. All participants filled out the demographic questionnaire and the prior knowledge test. Participants were randomly assigned to the gaze guidance ($n$ = 28) or no gaze guidance ($n$ = 28) condition; each group watched one video on declarative knowledge and one video on procedural knowledge. The order of the two videos was counterbalanced across participants. For each video lecture, eye-tracking data were collected in four phases. First, participants followed a dot on the screen with their eyes to finish calibration and validation. Second, they viewed their first assigned video lecture without pause. Third, they completed the learning performance test based on the knowledge learned from the video and completed the social presence scale. Fourth, they watched their other assigned video lecture, and data collection was repeated following the first three steps.

## 3 | RESULTS

Descriptive statistics and intercorrelations for all variables are shown in Tables 1, 2, and 3.

### 3.1 | Visual attention allocation

To test whether learners' visual attention allocation was guided by the instructor's gaze, we compared the total fixation duration on the PPT slides and the saccade counts from the area of the instructor to the PPT slides (H1).

First, analysing the fixation duration on the PPT slides, a mixed factor analysis of variance (ANOVA) revealed significant main effects of condition, $F(1, 56) = 34.48$, $p < 0.001$, $\eta^2 = 0.38$, and knowledge type, $F(1, 56) = 75.28$, $p < 0.001$, $\eta^2 = 0.57$. The main effect of condition indicated that participants viewing the video lectures with the instructor's gaze guidance spent more time looking at the PPT slides than participants viewing the video lectures without the instructor's gaze guidance. This result was in line with the hypothesis (H1). Furthermore, the main effect of knowledge type showed that participants spent more time looking at the learning content area while watching the procedural knowledge video lecture compared with the declarative knowledge video lecture.

Second, analysing the saccade counts between the area of the PPT slides and the area of the teacher image, a mixed factor ANOVA revealed significant main effects of condition, $F(1, 56) = 51.33$, $p < 0.001$, $\eta^2 = 0.48$, and knowledge type, $F(1, 56) = 200.68$, $p < 0.001$, $\eta^2 = 0.78$. Furthermore, the predicted interaction was significant, $F(1, 56) = 16.23$, $p < 0.001$, $\eta^2 = 0.23$. To further analyse the significant interaction, simple slope tests were conducted. The interaction effect (Figure 2) indicated that the participants viewing the video lectures with the instructor's gaze guidance switched attention fewer times between the area of the PPT slides and the area of the teacher image, and this difference was larger for the declarative knowledge videos, $Fs(1, 56) > 50.36$, $ps < 0.05$, than the procedural knowledge videos, $Fs(1, 56) > 14.67$, $ps < 0.05$.

This result was consistent with our hypothesis (H1) that in the video lecture under the instructor's gaze guidance condition, participants viewing the video lectures switched attention between the area of the instructor and the areas of the PPT slides less often than participants viewing the video lecture without the instructor's gaze guidance.

### 3.2 | Social presence

The hypothesis about social presence (H2) predicted that the participants viewing the video lectures with the instructor's gaze guidance would experience a higher level of social presence than the participants viewing the video lectures without the instructor's gaze guidance (H2). Mixed factor ANOVA showed a significant main effect of condition, $F(1, 56) = 50.07$, $p < 0.001$, $\eta^2 = 0.472$. This result was in line with our hypothesis (H2). Pairwise comparisons indicated that the instructor's gaze guidance in the video lecture increased participants' social presence, regardless of whether the video was on declarative knowledge or procedural knowledge. The results suggest that compared with the participants in the without eye gaze condition, the participants who saw the video lecture with the instructor's gaze guidance experienced more instruction interaction and felt more easily immersed into the learning environment.

### 3.3 | Learning performance

The hypothesis about learning performance (H3) was that participants viewing the video lectures with the instructor's gaze guidance would

**TABLE 1** Mean and standard deviation for social presence, visual attention allocation, and learning performance both declarative knowledge and procedural knowledge

| | | Video lecture without gaze guidance | | Video lecture with gaze guidance | |
|---|---|---|---|---|---|
| | Knowledge type | *M* | *SD* | *M* | *SD* |
| Social presence | Declarative | 3.24 | 0.36 | 3.85 | 0.56 |
| | Procedural | 3.15 | 0.51 | 3.91 | 0.48 |
| Fixation duration (AOI1) | Declarative | 150,200 | 14,046.08 | 169,300 | 14,982.59 |
| | Procedural | 165,700 | 7,626.13 | 180,400 | 11,804.33 |
| Saccade counts between two AOIs | Declarative | 20.38 | 5.77 | 11.13 | 3.78 |
| | Procedural | 7.88 | 4.02 | 4.33 | 2.96 |
| Learning performance | Declarative | 24.97 | 7.38 | 41.35 | 10.50 |
| | Procedural | 35.24 | 9.02 | 46.24 | 7.06 |

*Note*. AOI: area of interest.

**TABLE 2** Descriptive statistics and bivariate correlations for each outcome measure in declarative knowledge video lectures

| Measure | *M* | *SD* | 1 | 2 | 3 |
|---|---|---|---|---|---|
| 1. Social presence | 3.55 | 0.56 | – | | |
| 2. Fixation duration (AOI1) | 159,720 | 17,330.62 | 0.40** | – | |
| 3. Saccade counts between two AOIs | 15.84 | 6.66 | −0.51** | −0.63** | – |
| 4. Learning performance | 33.16 | 12.21 | 0.41** | 0.26 | −0.36** |

*Note*. AOI: area of interest. ** $p < 0.01$.

**TABLE 3** Descriptive statistics and bivariate correlations for each outcome measure in procedural knowledge video lectures

| Measure | *M* | *SD* | 1 | 2 | 3 |
|---|---|---|---|---|---|
| 1. Social presence | 3.52 | 0.63 | – | | |
| 2. Fixation duration (AOI1) | 173,042 | 12,366.06 | 0.49** | – | |
| 3. Saccade counts between two AOIs | 6.12 | 3.93 | −0.21 | −0.43** | – |
| 4. Learning performance | 40.74 | 9.76 | 0.45** | 0.40** | −0.39** |

*Note*. AOI: area of interest. ** $p < 0.01$.

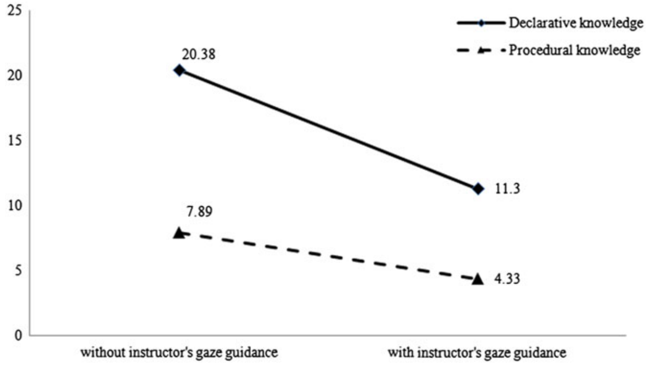


**FIGURE 2** Interaction effect of experimental conditions and knowledge type on the saccade counts between the area of the PowerPoint slides and the area of teacher image [Colour figure can be viewed at wileyonlinelibrary.com]

have higher learning scores than participants viewing the video lectures without the instructor's gaze guidance but only when learning procedural knowledge. Using the nature of the knowledge content as the independent variable, we conducted dependent sample *t* tests to test the extent to which learning performance was indeed promoted by the instructor's gaze guidance. The results showed that participants viewing the video lectures with the instructor's gaze guidance had significantly higher scores than those viewing the video lectures without the instructor's gaze guidance, regardless of whether they learned declarative knowledge, $t(56) = 6.87$, $p < 0.001$, $d = 0.68$, or procedural knowledge, $t(56) = 5.17$, $p < 0.001$, $d = 0.57$. The results partially supported our hypothesis (H3) and suggested that instructors' gaze guidance facilitated the learning of both types of knowledge.

### 3.4 | The moderation effects of social presence in the relation between the experimental condition and the learning performance

To test the moderating effect of social presence in the relation between the experimental condition (gaze guidance vs. no guidance) and the learning performance, two separate hierarchical multiple regression models were tested: one for the declarative video lecture and one for the procedural video lecture. As depicted in Table 4, social presence significantly moderated the relation between experimental condition and learning performance for procedural knowledge but not for declarative knowledge. To further analyse the significant interaction, simple slope tests were conducted. Learners with a high level of social presence showed better learning performance when they viewed the video lecture with gaze guidance than when they viewed the video lecture without gaze guidance (simple slope = 0.51, $t = 3.14$, $p < 0.01$). However, learners with a low level of social presence showed similar learning performance when they viewed the video lecture with gaze guidance and without gaze guidance (simple slope = 0.33, $t = 1.77$, $p > 0.05$). The results support our hypothesis (H4) and suggest that the instructor's gaze guidance has a stronger positive effect on learning when the learner has a high level of social presence.

## 4 | DISCUSSION

This study tested the effects of an instructor's gaze guidance in video lectures on learners' visual attention allocation, social presence, and learning performance, taking into account whether the content of the lecture was procedural or declarative knowledge. Results confirmed that the instructor's gaze guidance was beneficial for learning.

**TABLE 4** The moderating effects of social presence in the relation between the experimental condition and the learning performance

| | | | $\Delta R^2$ | *B* | *SE* | *β* |
|---|---|---|---|---|---|---|
| Declarative knowledge | Step 1 | | 0.46*** | | | |
| | | Experimental condition | | 15.60 | 2.87 | 0.64*** |
| | | Social presence | | 1.28 | 2.58 | 0.06 |
| | Step 2 | | 0.01 | | | |
| | | Experimental condition | | 15.13 | 2.93 | 0.63*** |
| | | Social presence | | 2.00 | 2.72 | 0.09 |
| | | Experimental condition × Social presence | | −0.20 | 0.24 | −0.09 |
| Procedural knowledge | Step 1 | | 0.34*** | | | |
| | | Experimental condition | | 9.04 | 2.68 | 0.47** |
| | | Social presence | | 2.59 | 2.16 | 0.17 |
| | Step 2 | | 0.05* | | | |
| | | Experimental condition | | 8.91 | 2.60 | 0.46** |
| | | Social presence | | 2.92 | 2.10 | 0.19 |
| | | Experimental condition × Social presence | | −0.36 | 0.17 | −0.22* |

*Note*. *β*: standardized regression coefficients; *B*: unstandardized regression coefficients; *SE*: standard error of the unstandardized regression coefficients. **p* < .05; ***p* < .01; ****p* < .001.

Meanwhile, learners viewing video lectures with the instructor's gaze guidance allocated more visual attention to the relevant areas of the PPT slides and perceived a higher level of social presence than those viewing video lectures without the instructor's gaze guidance. Overall, these results suggest that the instructor's gaze guidance is a powerful cue for learners to switch their visual attention from the instructor's face to the content area on the slides. In the remaining part of this section, we use Mayer's multimedia learning theory and social presence theory as frameworks for considering the relation between our findings and those of previous studies on video lectures, and we point to the strengths and limitations of the current study.

The data on learners' eye movements showed that the instructor's gaze guidance effectively directed learners' visual attention to the corresponding areas of the PPT slides, which supported our first hypothesis. In addition, when learners viewed the video with the instructor's eye gaze guidance, they were faster to switch gaze direction between the instructor and the PPT than when viewing the video in which the instructor did not change gaze direction (i.e., only looked into the camera). The results are consistent with Mayer's multimedia learning theory, which holds that the instructor's gaze direction could act as an attentional cue (Mayer, 2005). They are also consistent with the views of social cognition researchers who have described gaze guidance as one of the most salient ostensive signals in human communication for conveying communicative intent (Csibra & Gergely, 2009; Leong et al., 2017). Therefore, further studies could test whether the instructor's gaze guidance is superior to nonhuman cues in directing learners' attention.

Furthermore, the results related to social presence are in line with our second hypothesis. Learners who viewed video lectures with the instructor's gaze guidance perceived a higher level of social presence than those who viewed video lectures without the instructor's gaze guidance. This finding indirectly indicated that learners felt a sense of interaction with the instructor while learning with video lectures. Furthermore, during our informal discussions at the end of the experiment, the participants reported that the interactive function of the instructor's gaze guidance made them easily immerse themselves in the video lectures. Together, the results were consistent with social presence theory. According to this theory, social cues provided by the instructor (e.g., gestures, facial expression, and gaze guidance) are assumed to activate the learner's social response (e.g., smiling and tracking) and learning engagement and thus improve the learner's level of social presence (Dunsworth & Atkinson, 2007; F. X. Wang et al., 2018). These findings suggest that the instructor's gaze guidance might be a key social cue in video lectures.

Last but not least, the results of the learning performance test showed that the positive effects of the instructor's gaze guidance were present for video lectures on both declarative knowledge and procedural knowledge. One possible underlying mechanism is that the instructor's gaze guidance directed learners' attention. Another possible underlying mechanism could be that the instructor's gaze guidance improved learners' level of social presence. As a result, learners' attention allocation and high level of social presence may contribute to learning performance. Interestingly, we found that social presence had a moderating effect in the relation between the experimental condition (gaze guidance vs. no gaze guidance) and the learning procedural knowledge. Specifically, the instructor's gaze guidance in the procedural video lecture sharply enhanced the learning performance of the learners who experienced a high level of social presence but not of the learners who experienced a low level of social presence; if they reported a low level of social presence, participants' learning performance changed similarly in the two gaze guidance conditions.

The main reason for the moderating effect might have to do with the difference between learners' processing style. Learners with a high level of social presence may use a relationship-oriented processing style that focuses more on social interaction and mutual relationships, whereas learners with a low level of social presence might apply a task-oriented processing style with a focus on goal achievement (Cohen, Solomon, Maxfield, Pyszczynski, & Greenberg, 2004; Tabernero, Chambel, Curral, & Arana, 2009). In our study, the learners with a higher level of social presence were sensitive to social interactivity in the learning environment and tended to follow the teacher's gaze guidance to the key information. By contrast, the learners with a lower level of social presence focused on task learning and did not follow the teacher's gaze guidance, so learning performance was similar in the two conditions.

The study has several limitations. First, the conclusions about gaze guidance are based on one experiment. The learning environment differed from naturalistic learning settings, and learners' motivation might have differed in the two settings. Previous studies showed that the learning environment and learners' motivation affected their learning process and outcomes (Berland & Wilensky, 2015; Jeong & Gonzalez-Gomez, 2016). The laboratory setting in the current study might have influenced the results to a certain extent. Therefore, field studies and longitudinal studies are needed to test the beneficial effects of instructors' gaze guidance in naturalistic learning settings. Furthermore, the learning content was one topic from an education technology course, but even within this one topic, attention and learning varied depending on whether the topic was presented with a focus on declarative or procedural knowledge. This suggests that there might be great variability across topics, and the conclusions from this study may not generalize to other subjects.

To our knowledge, this is the first study to examine the underlying mechanism of the positive effect of instructors' gaze guidance in video lectures while taking into account the procedural versus declarative nature of the learning content. We have also extended the literature in this area by testing whether gaze guidance improves social presence and whether social presence moderates the benefits seen in other outcome variables. The results of this study add important information to the literature on video lectures and lead to a strong recommendation for educational practitioners: Instructors should provide gaze guidance in video lectures for better learning performance.

## ORCID

*Zhongling Pi* http://orcid.org/0000-0002-8776-6177

## REFERENCES

Alibali, M. W., Young, A. G., Crooks, N. M., Yeo, A., Wolfgram, M. S., Ledesma, I. M., & Knuth, E. J. (2013). Students learn more when their teacher has learned to gesture effectively. *Gesture*, *13*(2), 210–233.

Amadieu, F., Marine, C., & Laimay, C. (2011). The attention-guiding effect and cognitive load in the comprehension of animations. *Computers in Human Behavior*, *27*(1), 36–40.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*, 369–406.

Anderson, J. R. (1995). *Cognitive psychology and its implications* (4th ed.). New York, NY: WH Freeman.

Berland, M., & Wilensky, U. (2015). Comparing virtual and physical robotics environments for supporting complex systems and computational thinking. *Journal of Science Education & Technology*, *24*(5), 628–647.

Bos, N., Groeneveld, C., van Bruggen, J., & Brand-Gruwel, S. (2016). The use of recorded lectures in education and the impact on lecture attendance and exam performance. *British Journal of Educational Technology*, *47*(5), 906–917.

Cohen, F., Solomon, S., Maxfield, M., Pyszczynski, T., & Greenberg, J. (2004). Fatal attraction: The effects of mortality salience on evaluations of charismatic, task-oriented, and relationship-oriented leaders. *Psychological Science*, *15*(12), 846–851.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153.

De Koning, B. B., & Tabbers, H. K. (2013). Gestures in instructional animations: A helping hand to understanding non-human movements? *Applied Cognitive Psychology*, *27*(5), 683–689.

Dunsworth, Q., & Atkinson, R. K. (2007). Fostering multimedia learning of science: Exploring the role of an animated agent's image. *Computers & Education*, *49*(3), 677–690.

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.

Gunawardena, C. N. (1995). Social presence theory and implications for interaction and collaborative learning in computer conferences. *International Journal of Educational Telecommunications*, *1*, 147–166.

Hong, J., Pi, Z., & Yang, J. (2018). Learning declarative and procedural knowledge via video lectures: Cognitive load and learning effectiveness. *Innovations in Education and Teaching International*, *55*(1), 74–81.

Jeong, J. S., & Gonzalez-Gomez, D. (2016). Students' perceptions and emotions toward learning in a flipped general science classroom. *Journal of Science Education and Technology*, *25*(5), 747–758.

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.

Joksimovic, S., Gasevic, D., Kovanovic, V., Riecke, B. E., & Hatala, M. (2015). Social presence in online discussions as a process predictor of academic performance. *Journal of Computer Assisted Learning*, *31*(6), 638–654.

Kim, T., & Biocca, F. (1997). Telepresence via television: Two dimensions of telepresence may have different connections to memory and persuasion. *Journal of Computer-Mediated Communication* (Online), *3*(2).

Kizilcec, R. F., Papadopoulos, K., & Sritanyaratana, L. (2014, April). Showing face in video instruction: Effects on information retention, visual attention, and affect. *ACM Conference on Human Factors in Computing Systems (pp.2095-2102)*. ACM.

Leong, V., Byrne, E., Clackson, K., Georgieva, S., Lam, S., & Wass, S. (2017). Speaker gaze increases information coupling between infant and adult brains. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(50), 13290–13295.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In *Cambridge handbook of multimedia learning* (pp. 31–48). New York: Cambridge University Press.

Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology*, *93*(2), 390–397.

O'Bannon, B. W., Lubke, J. K., Beard, J. L., & Britt, V. G. (2011). Using podcasts to replace lecture: Effects on student achievement. *Computers & Education*, *57*(3), 1885–1892.

Ouwehand, K., Van Gog, T., & Paas, F. (2015). Effects of gestures on older adults' learning from video-based models. *Applied Cognitive Psychology*, *29*(1), 115–128.

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

Ozcelik, E., Arslan-Ari, I., & Cagiltay, K. (2010). Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior*, *26*(1), 110–117.

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.

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.

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.

Rueckert, L., Church, R. B., Avila, A., & Trejo, T. (2017). Gesture enhances learning of a complex statistical concept. *Cognitive Research Principles & Implications*, *2*(1), 2.

Tabernero, C., Chambel, M. J., Curral, L., & Arana, J. M. (2009). The role of task-oriented versus relationship-oriented leadership on normative contract and group performance. *Social Behavior & Personality an International Journal*, *37*(10), 1391–1404.

Traphagan, T., Kucsera, J. V., & Kishi, K. (2010). Impact of class lecture webcasting on attendance and learning. *Educational Technology Research & Development*, *58*(1), 19–37.

Wang, F. X., Li, W. J., Mayer, R. E., & Liu, H. S. (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.

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.

Wermeskerken, M., & 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.

Winslett, G. (2014). What counts as educational video? Working toward best practice alignment between video production approaches and outcomes. *Australasian Journal of Educational Technology*, *30*, 487–502.

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 & Implications*, *2*(1), 41.

**How to cite this article:** Wang H, Pi Z, Hu W. The instructor's gaze guidance in video lectures improves learning. *J Comput Assist Learn*. 2019;35:42–50. https://doi.org/10.1111/jcal.12309