American Psychological Association

Journal of Educational Psychology

0022-0663/19/$12.00

2019, Vol. 111, No. 8, 1382–1395
http://dx.doi.org/10.1037/edu0000352

Getting the Point: Which Kinds of Gestures by Pedagogical Agents Improve Multimedia Learning?

Wenjing Li and Fuxing Wang
Central China Normal University

Richard E. Mayer
University of California, Santa Barbara

Huashan Liu
Central China Normal University

Previous studies have shown that students learn better from an online lesson when a gesturing pedagogical agent is added (Mayer & DaPra, 2012; Wang, Li, Mayer, & Liu, 2018). The goal of this study is to pinpoint which aspect of a gesturing pedagogical agent causes an improvement in learning from an online lesson. College students learned about neural transmission in an online multimedia lesson that included a pedagogical agent who displayed specific pointing gestures (i.e., pointing to the specific component in the diagram being mentioned in the narration), general pointing gestures (i.e., pointing in the general direction of the diagram), nonpointing gestures (moving hands as beats, moving an arm up or down, or crossing two hands), or no gestures. An analysis of students' eye movements during learning showed that students in the specific-pointing group paid more attention to task-related elements than did students in the other groups (as indicated by fixation time and fixation count on the target area of interest). Students in the specific-pointing group also performed better than the other groups on retention and transfer tests administered immediately after the lesson and after a 1-week delay. The results show that an active ingredient in effective pedagogical agents is the use of specific pointing gestures. This work helps clarify the embodiment principle and image principle by isolating specific pointing (or deictic gestures) as a key feature that makes gesturing effective in multimedia lessons.

Educational Impact and Implications Statement

When students learn from multimedia lessons with onscreen pedagogical agents, does the type of gesturing by the agent affect student learning? In this study, students learned better and paid more attention to relevant material on the screen when the onscreen agent used specific pointing gestures during instruction rather than general pointing gestures, nonpointing gestures, or no gestures. Specifically pointing to relevant parts of the onscreen graphic while talking guides the learner's attention and leads to better learning.

Keywords: pedagogical agents, multimedia learning, gesture, pointing, deictic gestures

Consider an instructional scenario in which an instructor stands next to a screen that displays graphics on a scientific topic and orally explains what is on the screen. The present study examines the instructional impact of one feature in this situation—the kind of gesturing displayed by the instructor as she talks about what is on the screen in a lesson on neural transmission. In particular, the instructor in the present study is a human-like animated pedagogical agent, which allows us to systematically control the type and amount of gesturing. Our goal is not to examine educational technology per se, but rather to use it to study the broader issue of how the instructor's embodiment affects learning. In this case, we seek to better understand which kinds of gesturing affect learning outcomes and learning processes.

This research on the role of gesturing in instructional communications applies to instruction with pedagogical agents and also has implications for the design of face-to-face classroom instruction—particularly when an instructor is lecturing with a slideshow—and for the design of instructional video that could be used in Massive Open Online Course (MOOC), blended or flipped classrooms, or in informal online learning (e.g., a You-

This article was published Online First March 25, 2019.

Wenjing Li and Fuxing Wang, School of Psychology, Central China Normal University; Richard E. Mayer, Department of Psychological and Brain Sciences, University of California, Santa Barbara; Huashan Liu, School of Psychology, Central China Normal University.

Wenjing Li is now at the Academy of Psychology and Behavior, Tianjin Normal University.

This research was supported by the National Natural Science Foundation of China Grant 31771236 to Fuxing Wang.

Correspondence concerning this article should be addressed to Fuxing Wang, School of Psychology, Central China Normal University, 152 Luoyu Street, Wuhan, Hubei 430079, P. R. China. E-mail: fxwang@mail.ccnu.edu.cn

Tube video). The larger educational issue is how the instructor can guide the learner's attention to relevant parts of an instructional graphic during learning in a way that improves learning processes and outcomes, including in slideshow lectures (e.g., Clark & Mayer, 2016) and instructional video (Fiorella & Mayer, in press).

Research on instructor embodiment has shown that students learn better when the instructor gestures while orally explaining onscreen graphics rather than simple stands still (Mayer, 2014b; Mayer & DaPra, 2012; Wang, Li, Mayer, & Liu, 2018) or when the instructor draws the graphics on board as she talks rather than stands next to already completed drawings (Fiorella & Mayer, 2016). However, all forms of gesturing may not be equally effective, so an important next research step is to disentangle which aspect of the instructor's hand motions are crucial for causing an improvement in learning. The broader theoretical issue concerns whether gestures are helpful because they specifically guide the learner's attention to where to look in an accompanying graphic (as caused by specific pointing gestures [SPGs]) or because they remind the learner to look at the graphic (as caused by general pointing gestures [GPGs]), or because they build a social bond that fosters higher engagement (as caused by general nonpointing gestures [NPGs]). The social bond explanation comes from social agency theory (Mayer, 2014b; Mayer & DaPra, 2012), which posits that social cues such as human-like gesturing prime a social response in the learner based on seeing the pedagogical agent as a social partner, which leads to more effort to understand what the social partner is saying and therefore deeper learning outcomes. The present study helps us better tease apart which aspects of gesturing are responsible for guiding learning processes and causing learning outcomes. This study contributes more broadly to our understanding of how the instructor's embodiment affects student learning.

In particular, we compare the effectiveness of a pedagogical agent who engages in specific pointing gestures (i.e., the agent points to the component in the graphic when she mentions it), GPGs (i.e., the agent points in the general direction of the graphic when she mentions a component in the graphic), NPGs (i.e., the agent moves her hand at the same beat as her narration, crosses her two hands as she speaks, and puts her arm up or down as she speaks), and no gestures (NGs; i.e., the agent stands still as she speaks). This study seeks to disentangle which types of gesturing result in better learning on tests of learning outcome (based on immediate and delayed tests of retention and transfer), more visual attention paid to the relevant parts of the graphic (based on eye-tracking measures), and more positive ratings of the learning experience (based on self-report ratings). This issue has theoretical implications for how gesturing cues affect the learning process and practical implications for how to design effective onscreen agents.

Recent advances in computer technology allow instructional designers to add embodied animated pedagogical agents (APAs) displaying humanlike gestures, facial expression, body movement, and eye-gaze to an online lesson. Animated pedagogical agents are computer-generated characters who appear on the screen during a computer-based lesson and are intended to help students learn (Baylor & Ryu, 2003; Johnson & Lester, 2016; Veletsianos & Russell, 2014). The implementation of such highly embodied pedagogical agents is an attempt to increase instructional support and motivational elements in multimedia learning environments (Heidig & Clarebout, 2011). Although APAs have been widely used in online lessons for decades dating back to the 1990s (Johnson & Lester, 2016), evidence-based guidelines for how to design APAs to optimize learning are scarce. This study contributes more broadly to the guidelines for instructional practice focused on the role of instructor gesturing during lecturing.

According to the indexical hypothesis, language comprehension requires indexing (or connecting) abstract language (such as the instructor's lecture in the present study) to specific objects and situations (such as the graphic displayed on the screen in the present study) so that the affordances of the objects can constrain the combination of the ideas. Affordances arise from the interaction of an observer and an object (Glenberg & Robertson, 1999). In terms of the indexical hypothesis, gestures may be crucial for language comprehension, because they help the learner connect the spoken words with concrete elements. For example, Glenberg and Robertson (1999) found that students learned how to perform manual tasks better when the instructor pointed to or actually manipulated the key parts of an object as he orally mentioned them in explaining what to do. Thus, in the present study our focus is on the role of gesturing in helping learners in the indexing process of connecting words and components in graphics.

We begin with the idea that pedagogical agents who exhibit gesture may take several different forms. For example, they might accurately guide learners to a specific element in a graphic (e.g., Baylor & Kim, 2009) or just roughly to the entire material by their eye gaze, gesture, and body rotation (e.g., Dunsworth & Atkinson, 2007). They also might display casual movement (such as nodding; e.g., Baylor & Ryu, 2003) or just be static (e.g., Park, 2015). Although there is a growing body of research on embodied pedagogical agents, specific guidelines for designing such pedagogical agents to optimize learning are needed. Thus, the main objective of this study is to extend previous research on embodiment in pedagogical agents to explore how different kinds of gesturing movements affect multimedia learning.

Literature Review

One prominent question of interest to researchers and practitioners concerns how to design effective high-embodied pedagogical agents. There are some studies showing that students learned more deeply from a high-embodied agent who engaged in humanlike behaviors such as gesturing than a low-embodied agent who simply stood still (Baylor & Kim, 2009; Craig, Twyford, Irigoyen, & Zipp, 2015; Mayer & DaPra, 2012; Twyford & Craig, 2013; Wang et al., 2018). For example, Mayer and DaPra (2012) asked students to learn how solar cells work by watching a narrated presentation that included a low-embodied agent standing motionless or a high embodied agent displaying embodied cues such as human-like facial expression, eye gaze, and body movement. Three experiments consistently found that learners performed better on transfer tests when a human-voiced agent displayed embodied cues than when the agent did not.

Similarly, with the help of eye-tracking, Wang et al. (2018) asked students to learn the process of chemical synaptic transmission in a narrated presentation when the agent included embodied cues such as gesturing or not. The results showed that students who learned with an agent displaying embodied cues outperformed those who learned with a static agent on retention and transfer tests and focused their visual attention more on the relevant areas of the graphic as reflected by eye-tracking data.

These studies provide support for the embodiment principle: People learn better when on-screen agents display humanlike gesturing, movement, eye contact, and facial expression (Mayer, 2014b). In a review of research on embodiment in pedagogical agents, Mayer (2014b) reported that, in 11 out of 11 experimental comparisons, people performed better on transfer tests when they learned from a high-embodied agent than from a low-embodied agent, yielding a median effect size of $d = 0.36$. However, there are also some studies finding that students did not learn better from a narrated presentation when a low-embodied agent was changed into a high-embodied agent (Baylor & Ryu, 2003; Craig, Gholson, & Driscoll, 2002; Dirkin, Mishra, & Altermatt, 2005; Frechette & Moreno, 2010; Lusk & Atkinson, 2007). The possible explanation for inconsistent results may be that the embodiment principle is influenced by mitigating factors, that is, the voice of agent, the type of instructional material, the learner's prior knowledge, and so on. For example, Mayer and DaPra (2012) reported that the embodiment effect was found when the agent spoke in a human voice but not in a machine voice. Baylor and Kim (2009) found that the embodiment principle worked when students learned procedural instruction but not for attitudinal instruction. Overall, there may be other underlying factors affecting the embodiment effect that need to further explored, such as the type of gesture used by the agent.

Some studies have shown that gestures are an essential feature of human communication in general, because the act of communication involves integrating both what the speaker says and how the speaker gestures while speaking (Goldin-Meadow, 2003; McNeil, 2005). Research in educational communication has shown that gestures are beneficial to instruction and learning (Fiorella & Mayer, 2016; Macedonia, Müller, & Friederici, 2011; McGregor, 2008; McGregor, Rohlfing, & Marschner, 2009; Singer & Goldin-Meadow, 2005; Valenzeno, Alibali, & Klatzky, 2003). For example, Valenzeno et al. (2003) asked preschool children to view videotaped lessons about the concept of symmetry. They found that children who saw the lesson with gestures scored higher on the posttest than those who saw the lesson without gestures. The foregoing research results demonstrate the need for continued research on the role of embodiment in computer-based learning and particularly the effects of the pedagogical agent's gesturing on learning (Mayer, 2014b; Wang et al., 2018; Wang, Li, Xie, & Liu, 2017).

However, the kinds of gestures exhibited by highly embodied agents in previous studies were not always the same. In some studies, the pedagogical agents exhibited specific pointing gestures toward the task-related element on a graphic with their hand, eye gaze, and/or body rotation when the narration referred to it (Baylor & Kim, 2009; Johnson, Ozogul, Moreno, & Reisslein, 2013; Johnson, Ozogul, & Reisslein, 2015; Moreno, Reislein, & Ozogul, 2010; Twyford & Craig, 2013; Wang et al., 2018); in other studies, the pedagogical agents exhibited general gestures that guided learners to look in the direction of the presentation using their hand, eye gaze, and/or body rotation (Craig et al., 2002; Dunsworth & Atkinson, 2007; Lusk & Atkinson, 2007; Mayer & DaPra, 2012; Yung & Paas, 2015); in yet other studies, pedagogical engaged in casual NPGs without any specific visual guidance (Baylor & Ryu, 2003; Beege, Schneider, Nebel, & Rey, 2017; Shiban et al., 2015). Considering that different kinds of gestures may have different influence on learning (Gawne, Kelly, & Unger, 2009; Hostetter, 2011) and majority of previous studies are vague in the descriptions of the gestures displayed by pedagogical agents, it is worthwhile to disentangle the effects of different kinds of pedagogical agent gestures on instructional effectiveness in multimedia learning.

Theory and Predictions

The present study is informed by two research-based principles of multimedia design—the signaling (or cueing) principle and the embodiment principle. The *signaling principle* (also called the *cueing principle*) states that people learn better from multimedia lessons when the key material is highlighted (Mayer & Fiorella, 2014). Pointing gestures can be considered a form of visual signaling or cueing to the extent that they highlight where to look in a graphic (Wang et al., 2018). According to the cognitive theory of multimedia learning, signaling is effective because it reduces the learner's extraneous processing during learning, that is, cognitive processing that does not serve the instructional goal (Mayer, 2014a). Learners do have to scan around the illustrations to try to find the elements that the narration is mentioning if the onscreen agent points to the elements as she mentions them in her narration. Given the learner's limited capacity for processing material in working memory, this frees up cognitive capacity that can be used for essential processing (i.e., representing the words and graphics in working memory) and generative processing (i.e., organizing them and making connections between them). When corresponding spoken words and components in the graphics are in working memory at the same time, this facilitates making connections between them—that is, integrating the pictorial and verbal representation and prior knowledge—which is a crucial process in meaningful learning (Mayer, 2014a). Therefore, according to the signaling (or cueing) principle and the cognitive theory of multimedia learning from which it is derived, we make the following predictions:

> *Hypothesis 1a:* Students who learn from a pedagogical agent who engages in SPGs (SPG group) will perform better on tests of learning outcome than students who learn from a pedagogical agent who engages in GPGs (GPG group), NPGs (NPG group), or NGs (NG group).

> *Hypothesis 2a:* Students who learn from a pedagogical agent who engages in SPGs (SPG group) will visually attend to the relevant portion of the graphic more (based on eye-tracking measures) than students who learn from a pedagogical agent who engages in GPGs (GPG group), NPGs (NPG group), or NGs (NG group).

> *Hypothesis 3a:* Students who learn from a pedagogical agent who engages in SPGs (SPG group) will report more positive ratings of their learning experience than students who learn from a pedagogical agent who engages in GPGs (GPG group), NPGs (NPG group), or NGs (NG group).

In contrast, the embodiment principle is that people learn better from multimedia lessons when on-screen agents display humanlike gesture, movement, eye contact, and facial expression (Mayer, 2014a). In particular, the present study compares three types of humanlike gesturing—SPGs, GPGs, and NPGs against NGs. According to social agency theory, gestures can serve as a kind of social cue that primes a social stance in learners causing them to work harder to make sense of presented material (i.e., engage in the cognitive processes of organizing and integrating), which finally leads to better learning outcomes (Fiorella & Mayer, 2016; Mayer, 2014a; Mayer & DaPra, 2012; Wang et al., 2018). According to the embodiment principle and social agency theory, all natural gestures should be equally effective as social cues and be better than no gestures on measures of learning outcome and learning process. In terms of the embodiment principle and social agency theory, we can make the following predictions.

> *Hypothesis 1b:* Students who learn from a pedagogical agent who engages in SPGs (SPG group), GPGs (GPG group), or NPGs (NPG group) will perform better on tests of learning outcome than students who learn from a pedagogical agent who engages in NGs (NG group).

> *Hypothesis 2b:* Students who learn from a pedagogical agent who engages in SPGs (SPG group), GPGs (GPG group), and NPGs (NPG group) will visually attend to the relevant portion of the graphic (based on eye-tracking measures) more than students who learn from a pedagogical agent who engages in NGs (NG group). If gesturing in general causes learners to work harder to make sense of the material, the elements on the slide should get more attention during learning when the pedagogical agent gestures.

> *Hypothesis 3b:* Students who learn from a pedagogical agent who engages in SPGs (SPG group), GPGs (GPG group), and NPGs (NPG group) will report more positive ratings of their learning experience than students who learn from a pedagogical agent who engages in NGs (NG group).

Eye-tracking methodology offers a unique way to examine the visual attention of learners during the learning process (Holmqvist et al., 2011; Mayer, 2010) and has contributed to research on computer-based instruction with onscreen instructors and graphics (Stull, Fiorella, & Mayer, 2018; Wang et al., 2018). Eye-tracking data are particularly helpful in addressing hypotheses 2a and 2b, which are instrumental in understanding how gestures can direct visual attention.

Method

Participants and Design

The participants were 123 undergraduates recruited from a university in central China via advertisements. Their mean age was 20.5 years ($SD = 2.3$) and 105 of them were women. The experiment used a one-factor between-subjects design with 32 participants in the SPG group, 30 in the GPG group, 31 in the NPG group, and 30 in the NG group. All participants had normal or corrected-to-normal vision, and Chinese was their native language. They were majoring in psychology ($n = 50$), education ($n = 10$), English ($n = 9$), biology ($n = 9$), mathematics ($n = 7$), physics ($n = 6$), history ($n = 5$), chemistry ($n = 4$), geography ($n = 4$), law ($n = 3$), computer ($n = 3$), economics ($n = 3$), Chinese ($n = 3$), management ($n = 2$), art ($n = 2$), music ($n = 2$), and politics ($n = 1$). There was no significant difference among the four groups on prior knowledge based on a pretest, $F(3, 119) = 1.39$, $p > .05$; mean age, $F(3, 119) = 1.10$, $p > .05$; and proportion of men and women, $\chi^2(3) = 0.24$, $p > .05$.

Procedure

Participants were tested individually in a lab setting. First, participants were randomly assigned to one of the four gesture conditions (SPG, GPG, NPG, or NG group) and completed the prequestionnaire at their own pace. Second, the experimenter asked participants to sit in front of the eye-tracking monitor and carried out the calibration process. Third, participants read the instructions and once they understood them, the multimedia lesson was presented via the computer system, which lasted 128 s. Following the presentation, participants first completed the postquestionnaire (consisting of four subjective rating questions and the Animated Persona Instrument [API]) and then the immediate tests (consisting of the retention test and transfer test) at their own pace. The entire session took about 35 min. After 1 week, participants came back and completed the delayed tests individually in the lab setting at their own pace. We adhered to guidelines for ethical treatment of human subjects and obtained institutional review board approval.

Materials

The instructional materials consisted of four versions of a computerized multimedia lesson that described how chemical signals were transmitted in the nervous system and the roles of action potential, calcium ions, sodium ions, and neurotransmitters in the transmission process. As illustrated in Figure 1, all versions included an illustration that showed the parts of neurons involved in synaptic transmission and an animated female agent standing to the left of the illustration. All versions had the same narrated script recorded by a young woman, lasting approximately 128 s and containing approximately 524 words. In the SPG version the pedagogical agent used posture, eye gaze, and pointing gesture (with a handheld pointer) to direct attention to the relevant parts of the illustration as they were discussed in the narration. Overall, 13 idea units were signaled by SPGs, which constitute the important elements from the 22 idea units that were presented. The signaled idea units were 3, 4, 5, 6, 7, 8, 9, 11, 13, 16, 18, 20, and 22 and are shown in Appendix A. The GPG version was the same as the SPG version, except the agent pointed toward the entire illustration as the relevant elements of the illustration were discussed in the narration. In the NPG version, the agent displayed casual gestures with her hands and arms, such as moving her hands in beat with the narration, putting her right or left arm up or down, or crossing her two hands. There was no pointing toward the relevant parts of the illustration or even toward the illustration in general. In the NG version, the agent was static without any gesturing. All multimedia lessons were created using Flash CS6 software with the screen size of 1,680 pixels × 1,050 pixels. All materials were in Chinese.

The prequestionnaire solicited demographic information (such as gender, age, and major) and included 10 multiple-choice questions about chemical synaptic transmission and four subjective rating statements that were intended to measure learners' prior

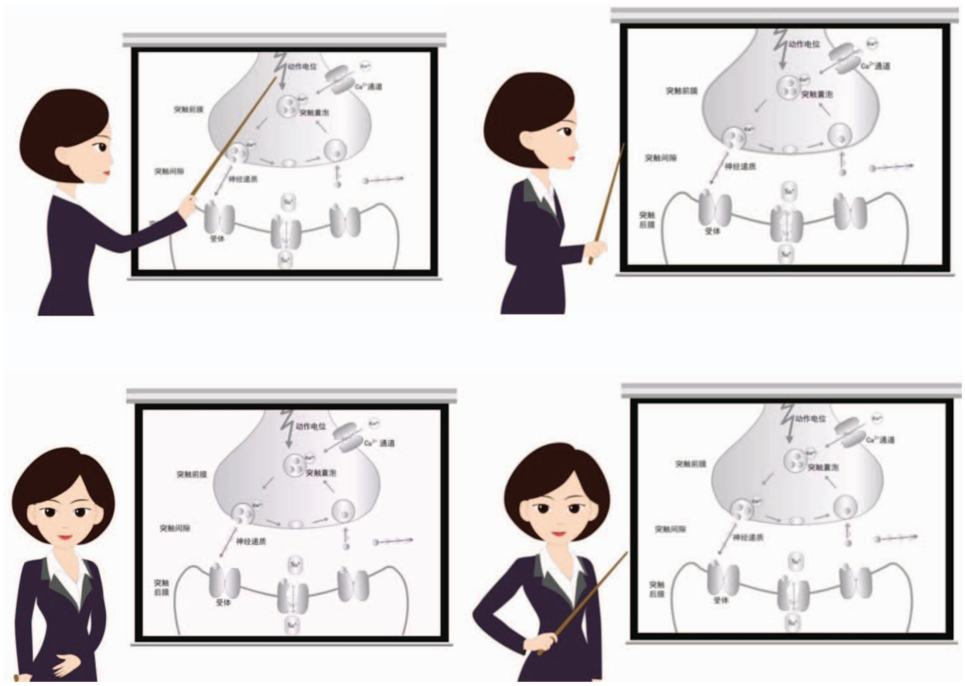

Figure 1. Example frames from four animations. Upper left panel: specific pointing gesture version; upper right panel: general pointing gesture version; lower left panel: nonpointing gesture version; lower right panel: no gesture version. See the online article for the color version of this figure.

acknowledge more comprehensively. An example question was "Where are neurotransmitters stored before they are released?" Each question had four options, and only one was the correct answer. Two points were awarded for each correct answer. An example rating statement is "How much do you know about chemical synapses?" The participants needed to respond on a five-point scale ranging from 0 (*very little*) to 4 (*very much*). A prior knowledge score was computed by summing the number of points on the multiple-choice items with the number of points on the rating items, yielding a score that could range from 0 to 31. The correlation coefficient between the objective and subjective knowledge scores was 0.52 ($p < .001$). The Cronbach's alpha for the prior knowledge items was 0.71.

The postquestionnaire included four subjective questions and the API. The four subjective questions were intended to assess learning perceptions concerning mental effort ("How much effort did you put in learning process?"), difficulty ("How difficult was the material you just learned?"), interest ("How interesting was the learning material?"), and motivation ("How much would you like to learn other learning materials in this way?"). These items were rated on a nine-point scale ranging from 1 (*very little*) to 9 (*very much*).

The API was used to assess the student's sense of social presence with the agent, which consisted of 25 items and clustered into four subscales measuring how well the agent facilitated learning, how human-like the agent seemed, how credible the agent seemed, and how engaging the agent was (Baylor & Ryu, 2003; Mayer & DaPra, 2012; Ryu & Baylor, 2005). Cronbach's alpha metrics for the API subscales were 0.95 for facilitated learning, 0.93 for human-like, 0.87 for credible, and 0.85 for engaging, which indicate suitable internal reliability.

The learning outcome tests consisted of a retention test and a transfer test. The retention test required participants to write down the process of chemical synaptic transmission in detail according to what they had learned. One point was awarded for each key point. A maximum of 22 points could be achieved. The 22 idea units are shown in Appendix A. The transfer test consisted of five open-ended questions (e.g., "What factors can affect the process of chemical synaptic transmission?"). Three acceptable answers for this question were the amplitude and schedule of action potential, concentration of Ca^{2+}, and the inactivation of neurotransmitters. The entire set of transfer questions is shown in Appendix B. In order to answer the transfer tests, learners must remember all 22 idea units and apply what they learned to solve novel problems. There are a total of 17 acceptable idea units on the transfer test. One point was awarded for each acceptable idea in the answer, thus the total score was 17 points. Two raters scored the retention and transfer tests independently and the average of them was the learner's final retention and transfer score. Interrater reliability of the immediate retention and transfer tests was 0.99 and 0.98, respectively (based on Pearson correlation, $ps < .001$). The delayed tests were same as the immediate tests except that the order of items was changed. The interrater reliability of the delayed retention and transfer tests both was 0.99 for both (based on Pearson correlation, $ps < .001$). The Cronbach's alpha for the transfer test was 0.76.

Apparatus

A SMI RED 250 Desktop eye-tracker (SensoMotoric Instruments, Teltow, Germany) was used to record the eye movement data. The eye tracker operates binocularly at a sampling rate of 250

Hz and has a spatial resolution of less than 0.1°. The computer screen for displaying the animation was positioned 65 cm from the participant with 680-pixels × 1,050-pixels resolution. The fixation filtering threshold was set at 100 ms.

Results

We compared the four groups on measures of learning outcome (immediate retention, immediate transfer, delayed retention, delayed transfer), measures of eye movements during learning (fixation duration and number of fixations on the target area of interest), and measures of learning perceptions (ratings on the API, effort, difficulty, interest, and motivation). For each set of measures as dependent variables, we conducted a one-way analysis of variance with group as the between-subjects factor. We used a significance level of 0.05 for all analyses and post hoc comparisons were Bonferroni-corrected. According to Cohen (1988), partial eta-squared and Cohen's *d* can be used as measures of effect size, with $\eta_p^2 = 0.01, 0.06, 0.14$ and $d = 0.20, 0.50, 0.80$ corresponding to small, medium, and large effects, respectively.

Learning Outcomes

Do different kinds of gesture affect learning outcomes on immediate tests? According to the signaling (or cueing) principle, students who view an agent who exhibits specific pointing gestures during instruction (SPG group) should learn better than students in the other groups (SPG, GPG, NPG groups). In contrast, according to the embodiment principle, students who view an agent who exhibits human-like gestures during instruction (SPG, GPG, NPG groups) should learn better than students who view an agent who does not move (NG group).

Table 1 shows the mean score (and standard deviation) for each group on the each of four tests. For the immediate retention test, the four groups differed significantly, $F(3, 119) = 9.54$, $p < .001$, $\eta_p^2 = .19$. Post hoc tests showed that the SPG group outperformed the GPG group ($d = 1.24$), NPG group ($d = 0.90$), and NG group ($d = 1.14$), which did not differ significantly from each other. For the immediate transfer test, the four groups differed significantly, $F(3, 119) = 9.66$, $p < .001$, $\eta_p^2 = .20$. Post hoc tests showed that the SPG group outperformed the GPG group ($d = 0.87$), NPG group ($d = 0.96$), and NG group ($d = 0.81$), which did not differ significantly from each other. The results on the immediate tests show that students learn best with an agent who exhibits specific gestures during instruction, and thereby the results are most consistent with the signaling (or cueing) principle.

Do different kinds of gestures affect learning outcomes on delayed tests? For the delayed retention test, the four groups differed significantly, $F(3, 119) = 12.28$, $p < .001$, $\eta_p^2 = .24$. Post hoc tests showed that the SPG group outperformed the GPG group ($d = 1.40$), NPG group ($d = 1.07$), and NG group ($d = 1.14$), which did not differ significantly from each other. For the delayed transfer test, the four groups differed significantly, $F(3, 119) = 11.05$, $p < .001$, $\eta_p^2 = .22$. Post hoc tests showed that the SPG group outperformed the GPG group ($d = 1.15$), NPG group ($d = 1.36$), and NG group ($d = 1.14$), which did not differ significantly from each other.

Overall, we conclude that delayed learning outcomes show the same pattern of results as the immediate learning outcomes, with the best learning outcome for the group that received a pedagogical agent who displayed specific pointing gestures. These results are consistent with the predictions of the signaling (or cueing) principle and suggest a boundary condition for the embodiment principle in which certain kinds of gesturing (i.e., specific pointing gestures) are more effective than others in promoting learning. These results can be understood from the perspective of the adjunct question paradigm, which demonstrated the benefits from taking the test immediately, reflected in substantial scores when the tests were given a day or even a week later (Anderson & Biddle, 1975).

Eye Movements During Learning

It is important for learners to attend to the information that is referred by narration in a timely manner because of the transience of information. By looking at the appropriate element in an illustration while the agent is talking about it, learners can make connections between corresponding words and graphics, which is a fundamental cognitive process in meaningful learning according to the cognitive theory of multimedia learning (Mayer, 2009, 2014a). To foster this integration process, learners may benefit from cues that can guide their attention to the right place at the

Table 1
Mean Scores and Standard Deviations on Learning Performance and Eye-Tracking Measures for Four Groups

Dependent variable	SPG *M*	SPG *SD*	GPG *M*	GPG *SD*	NPG *M*	NPG *SD*	NG *M*	NG *SD*	*F*	*p*	η_p^2
Prior knowledge	18.09	4.22	15.27	6.02	16.06	6.53	16.73	5.67	1.39	.248	.03
Immediate retention	7.00	2.74	3.89	2.37	4.41	2.99	4.08	2.39	9.54	<.001	.19
Immediate transfer	6.06	2.44	3.85	2.64	3.95	1.94	4.05	2.53	9.66	<.001	.20
Delayed retention	6.35	2.68	3.06	1.96	3.61	2.45	3.36	2.58	12.28	<.001	.24
Delayed transfer	6.95	2.25	4.08	2.73	3.94	2.19	4.18	2.58	11.05	<.001	.22
Fixation time (ms)	9642	3245	6018	2767	6952	3273	7459	2774	8.05	<.001	.17
Fixation count	29.47	9.21	19.23	8.63	21.77	9.09	23.00	8.05	7.75	<.001	.16
Revisits	11.28	3.66	7.00	4.05	7.74	4.10	8.97	3.42	7.44	<.001	.16
First fixation duration (ms)	221	82	165	71	182	72	173	73	3.47	.018	.08
Average fixation (ms)	232	93	159	52	186	67	174	62	6.13	.001	.13

Note. SPG = specific pointing gesture group; GPG = general pointing gesture group; NPG = non-pointing gesture group; NG = no gesture group; η_p^2 = partial eta-square.

right time. Therefore, in order to further assess the effectiveness of the pedagogical agent's gestures, we created 13 temporary Areas of Interest (AOIs) corresponding to each cued component in the illustration to find out whether the component was fixated at the time 5s after it was verbally evoked in the narration (Boucheix, Lowe, Putri, & Groff, 2013; Wang et al., 2018), as shown in Figure 2. We applied this time-locked analysis to each of the following five eye-tracking measures:

fixation time—the number of seconds the student fixates on relevant elements (e.g., the elements being described in the narrative),

fixation count—the number of times the student fixates on the relevant elements,

average fixation—sum of fixation time on relevant elements divided by number of fixations on relevant elements,

first fixation duration—the number of seconds the student first fixates on relevant elements, and

revisits—the number of times the student saccades to the relevant elements from outside after the first visit.

Table 1 shows the mean score (and standard deviation) for each of the four groups on the five eye-tracking measures. As with the learning performance, we conducted one-way analyses of variance for each eye-tracking measure. We used a significance level of 0.05 for all analyses and post hoc comparisons were Bonferroni-corrected. In short, we wanted to determine whether the groups differed in the degree to which they looked at the relevant portion of the illustration (i.e., corresponding to the component being mentioned in the narration), based on each of five eye-tracking measures. According to the signaling (or cueing) principle, students who view an agent who exhibits specific pointing gestures during instruction (SPG group) should focus more on relevant AOIs during the lesson than students in other groups (SPG, GPG, NPG groups). In contrast, according to the embodiment principle, students who view an agent who exhibits human-like gestures during instruction (SPG, GPG, NPG groups) should focus more on relevant AOIs during the lesson than students who view an agent who does not move (NG group).

For fixation time, the groups differed significantly, $F(3, 119) = 8.05$, $p < .001$, $\eta_p^2 = .17$. Post hoc tests showed that the SPG group had significantly longer fixation time on the relevant part of the illustration than the GPG ($d = 1.20$), NPG ($d = 0.83$) and NG groups ($d = 0.72$), which did not differ significantly from each other.

For fixation count, the groups differed significantly, $F(3, 119) = 7.75$, $p < .001$, $\eta_p^2 = .16$. Post hoc tests showed that the SPG group had a significantly higher fixation count than the GPG ($d = 1.15$), NPG ($d = 0.84$), and NG groups ($d = 0.75$), which did not differ significantly from each other.

For average fixation, the groups differed significantly, $F(3, 119) = 6.13$, $p < .01$, $\eta_p^2 = .13$. Post hoc tests showed that the SPG group had a significantly longer average fixation on the relevant part of the illustration than the GPG ($d = 0.97$), NPG ($p = .078$, $d = 0.57$), and NG groups ($d = 0.73$), which did not differ significantly from each other.

For first fixation time, the groups differed significantly, $F(3, 119) = 3.47$, $p < .05$, $\eta_p^2 = .08$. Post hoc tests showed that the SPG group had a significantly longer first fixation duration on the relevant part of the illustration than the GPG ($d = 0.74$) and NG groups ($p = .077$, $d = 0.62$), which did not differ significantly from each other.

For revisits, the groups differed significantly, $F(3, 119) = 7.42$, $p < .001$, $\eta_p^2 = .16$. Post hoc tests showed that the SPG group had

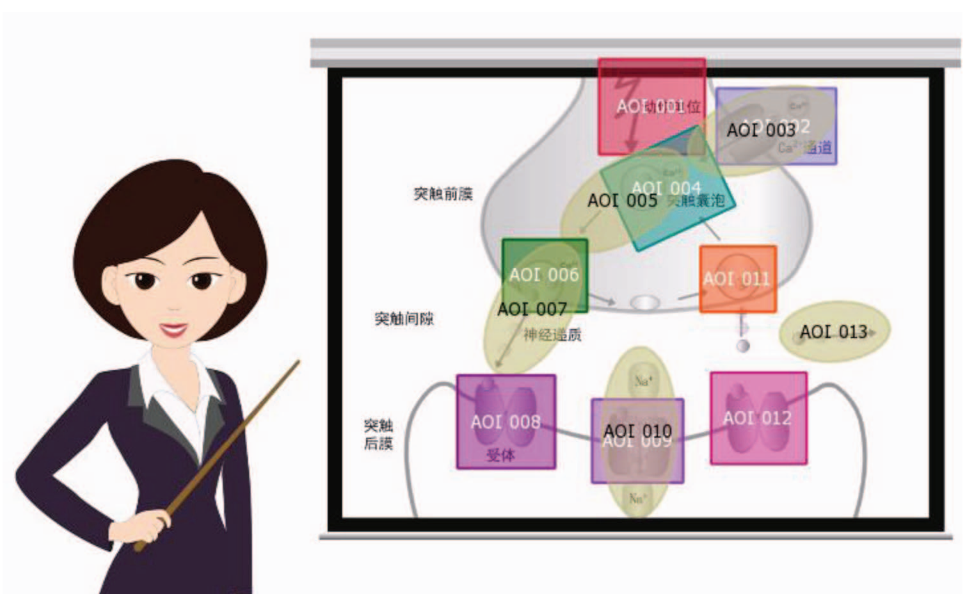

Figure 2. Thirteen temporary Areas of Interest (AOIs) for four groups. See the online article for the color version of this figure.

significantly more revisits to relevant areas than the GPG (d = 1.11) and NPG groups (d = 0.91), which did not differ significantly from each other.

We conclude that students who learned with a pedagogical agent using specific pointing gestures directed more of their visual attention to the relevant portion of the illustration corresponding to the part being mentioned in the narration than did students in each of the other groups. In short, the agent's pointing gesture in the SPG group was successful in guiding the learner's visual attention to relevant portions of the illustration and thereby facilitate the learner's ability to integrate corresponding words and components in the illustration. This finding confirms the signaling (or cueing) principle by showing that specific pointing is an effective way to direct the learner's visual attention.

Perceptions of Learning

The foregoing analyses of learning outcomes and visual attention during learning suggest that all forms of gesturing are not equally effective in promoting learning and guiding visual attention. A third source of information concerning possible differences among the groups comes from learners' subjective ratings of their learning experience. Table 2 shows the mean rating (and standard deviation) on the each of subjective questions, with higher scores indicating higher agreement. As with the learning performance and eye-tracking measures, we conducted one-way analyses of variance for each rating measure. We used a significance level of 0.05 for all analyses and post hoc comparisons were Bonferroni-corrected. For mental effort rating and the difficulty rating, there was no significant difference among the four groups ($F < 1$, $p >$ .05). For learning interest, the four groups differed significantly, $F(3, 119) = 3.06$, $p < .05$, $\eta_p^2 = .07$, and post hoc tests showed that the NPG group scored higher than the GPG group (p = .068, d = 0.66). For learning motivation, the four groups differed significantly, $F(3, 119) = 4.53$, $p < .01$, $\eta_p^2 = .10$, and post hoc tests showed that the SPG group (d = 0.80) and NPG group (d = 0.79) were higher than the GPG group.

For the API, the four groups differed significantly on one of the four dimensions: agent was credible, $F(3, 119) = 3.40$, $p < .05$, $\eta_p^2 = .08$, and post hoc tests showed that the SPG group (d = 0.75) gave higher ratings than the GPG group. This finding suggests that instructors who are more precise in how they gesture are considered to be more credible. There was no significant difference among the four groups on the other three dimensions: agent facilitated learning, $F(3, 119) = 2.18$, $p > .05$; agent was engaging, $F(3, 119) = 1.96$, $p > .05$; agent was human-like, $F(3, 119) = 1.79$, $p > .05$.

Overall, there is a pattern in which students perceive the lesson more positively (in terms of motivation and agent credibility) when the onscreen agent uses specific pointing gestures rather than general pointing gestures. This pattern is consistent with the signaling (or cueing) principle and suggests some boundary conditions for the embodiment principle in which certain kinds of gestures are appreciated more than others.

However, an unexpected result is that students learning from an agent with NPGs reported being more motivated than those who learned from an agent displaying general pointing gestures, which is not consistent with the signaling principle or the embodiment principle. One explanation could be derived from the coherence principle that people learn better when extraneous material is excluded rather than included (Mayer, 2009). The agent with general pointing gestures used her hand to guide learner's attention to the graphic rather than to the specific component in the graphic when she mentioned it, which is not directly beneficial for understanding the cause-and-effect explanation of neural transmission and may even distract from learning to some extent. The agent with NPGs is natural and far away from the graphic and therefore may require less attention from learners. Thus, students may have more positive perceptions of the agent with NPGs than the agent employing general pointing gestures, which may be seen as unhelpful.

Supplemental Analyses

The foregoing analyses found that students fixated more and longer on the important elements and perceived the lesson more positively when the agent used SPGs rather than GPGs. However, there remains the question of whether more fixations on elements and positive learning ratings are related to better learning outcomes, especially the relationship between distribution of attention and learning outcomes. The existing literature is inconclusive. For example, Wang et al. (2018) found that longer fixation time on cued elements was related to better retention and transfer performance. However, Ozcelik, Arslan-Ari, and Cagiltaym (2010) found that longer mean fixation duration on relevant information

Table 2
Mean Scores and Standard Deviations on Questionnaire Ratings for Four Groups

Dependent variable	SPG		GPG		NPG		NG				
	M	SD	M	SD	M	SD	M	SD	F	p	η_p^2
Mental effort	6.91	1.35	6.97	1.38	6.97	1.08	7.17	1.12	.26	.857	.01
Difficulty of materials	6.84	1.55	7.23	1.22	7.65	1.17	7.33	1.37	1.16	.327	.03
Learning interest	5.69	1.64	4.67	1.83	5.77	1.48	5.03	1.77	3.06	.031	.07
Learning motivation	6.06	1.92	4.50	1.96	5.90	1.56	5.37	1.87	4.53	.005	.10
Agent facilitated learning	29.09	10.00	23.73	7.23	26.97	8.10	28.30	9.69	2.18	.094	.05
Agent was credible	17.22	4.48	14.00	4.08	15.03	3.94	16.33	4.55	3.40	.020	.08
Agent was human-like	12.16	3.91	9.93	4.74	10.97	3.92	10.40	3.35	1.79	.153	.04
Agent was engaging	14.28	4.05	12.07	3.76	13.06	4.16	12.37	3.67	1.97	.123	.05

Note. SPG = specific pointing gesture group; GPG = general pointing gesture group; NPG = non-pointing gesture group; NG = no gesture group; η_p^2 = partial eta-square.

was related to worse transfer score. Therefore, we conducted supplemental analyses to begin to address this issue.

How do learning performance measures correlate with eye-tracking measures? In order to further explore whether learners' better learning outcomes was correlated with more effective attention guidance during learning, a correlation analysis was conducted among the four measures of learning outcome (i.e., immediate retention test score, immediate transfer test score, delayed retention test score, and delayed transfer test score) and five eye-tracking measures (i.e., fixation time, fixation count, average fixation, first fixation duration, and revisits. Table 3 shows that all 20 of these key correlations between a learning outcome measure and an eye-tracking measure are in the predicted direction, and 16 of the 20 correlations are statistically significant at the $p < .05$ level. What's more, fixation time, fixation count, average fixation and revisits on the important elements were all positively related to immediate retention score, immediate transfer score, delayed retention score, and delayed transfer score. This pattern of results is consistent with the idea that efficient visual processing during learning was related to superior scores on tests of learning outcome.

How do learning performance measures correlate with postquestionnaire ratings? In order to further explore whether learners' better learning outcomes was correlated with better learning perception during learning, a correlation analysis was conducted among the four measures of learning outcome (i.e., immediate retention test score, immediate transfer test score, delayed retention test score, and delayed transfer test score) and postquestionnaire scores (learning interest, learning motivation, and credibility of agent subscale). Table 4 shows that learning interest, learning motivation, and credibility of agent subscale were all positively related to immediate retention score, immediate transfer score, delayed retention score, and delayed transfer score. This pattern of results is consistent with the idea that superior scores on tests of learning outcome were related to positive subjective perceptions.

Discussion

Empirical Contributions

Although gesturing movements of pedagogical agents may play an important role in learning (De Koning & Tabbers, 2013; Johnson et al., 2015; Twyford & Craig, 2013), the issue of how different kinds of gestures influence learning needs to be clearly addressed. Through the current study, we found three major findings favoring the use of specific pointing gestures by pedagogical agents in an online multimedia lesson. First, consistent with hypothesis 1a, students learned better (as indicated by performance on immediate retention, immediate transfer, delayed retention, and delayed transfer tests) with a pedagogical agent who used SPGs rather than GPGs, NPGs, or NGs, which did not differ significantly from each other. Second, consistent with Hypothesis 2a, students attended to the relevant part of the illustration more often (as indicated by fixation time, fixation count, and average fixation) with a pedagogical agent who employed SPGs rather than GPGs, NPGs, or NGs, which did not differ significantly from each other. In this study, eye-tracking methodology made a unique contribution concerning learners' visual attention learning, demonstrating the overall value of eye-tracking data in learning research (Holmqvist et al., 2011; Mayer, 2010). Third, consistent with Hypothesis 3a, students rated the lesson more favorably (on motivation and credibility of the agent) if they learned from a pedagogical agent who exhibited specific pointing gestures rather than GPGs. Overall, there is strong evidence across multiple dependent measures, that pedagogical agents add more value to a multimedia lesson when they engage in SPGs—that is, point to the part of the illustration being mentioned in the narration—rather than other forms of gesture or no gesture at all.

Although some previous studies have indicated that adding embodiment cues to an agent (i.e., eye gaze and gestures) was

Table 4
Correlations Among Learning Outcome Scores and Postquestionnaire Scores

Measure	1	2	3	4	5	6	7
1. Immediate retention test	—						
2. Immediate transfer test	.70***	—					
3. Delayed retention test	.83***	.74***	—				
4. Delayed transfer test	.71***	.90***	.73***	—			
5. Learning interest	.22*	.27**	.31***	.26**	—		
6. Learning motivation	.16	.28**	.26**	.28**	.84***	—	
7. Agent was credible	.06	.20*	.14	.24**	.28**	.37***	—

* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 3
Correlations Among Learning Outcome Scores and Eye-Tracking Measures

Measure	1	2	3	4	5	6	7	8	9
1. Immediate retention	—								
2. Immediate transfer	.70***	—							
3. Delayed retention	.83***	.74***	—						
4. Delayed transfer	.71***	.90***	.73***	—					
5. Fixation time (ms)	.20*	.26**	.27**	.25**	—				
6. Fixation count	.24**	.29**	.28**	.29**	.76***	—			
7. Average fixation (ms)	.23*	.28**	.24**	.25**	.74***	.34***	—		
8. First fixation duration (ms)	.12	.17	.12	.17	.67***	.28**	.84***	—	
9. Revisits	.24**	.26**	.26**	.29**	.61***	.87***	.23*	.18*	—

* $p < .05$. ** $p < .01$. *** $p < .001$.

beneficial to learning (Mayer & DaPra, 2012; Wang et al., 2018), other studies found that adding embodiment cues to an agent was not beneficial to learning (i.e., Frechette & Moreno, 2010; Lusk & Atkinson, 2007). The results of the present study may provide a possible answer to the apparently inconsistent results in that some kinds of gestures (i.e., pointing gestures) are more effective than others.

Theoretical Implications

How does gesturing by pedagogical agents affect student learning from multimedia lessons? According to the signaling hypothesis, gestures can serve to highlight relevant parts of an onscreen graphic—such as an animation, video, or illustration. In particular, SPGs can guide the learner's attention to the part of the graphic that the agent is talking about. This allows for corresponding verbal and pictorial representations to be held in working memory at the same time so the learner can make connections between them—a cognitive process called *integrating* in the cognitive theory of multimedia learning (Mayer, 2009, 2014a). Integrating is considered to be a crucial step in constructing a meaningful learning outcome that can support performance on tests of learning outcome. According to the signaling hypotheses, gestures by pedagogical agents are effective to the extent that they signal where to look in the graphic, so we can predict that the specific pointing gesture group will outperform the other three groups on posttest measures of learning outcome, eye-tracking measures of attending to the relevant portion of the graphic, and positive self-report measures of the learning experience. The results are consistent with these predictions and the signaling hypothesis from which they are derived.

According to the embodiment hypothesis, gestures can serve as social cues that increase the learner's social connection with the onscreen agent and thereby cause the learner to exert more effort to make sense of the material. Accordingly, learning with a pedagogical agent who displays any kind of humanlike gesturing (i.e., SPGs, GPGs, and NPGs) should result in better learning outcomes, more visual attention to relevant portions of the graph and higher ratings than learning with a pedagogical agent who does not gesture. The results are not consistent with this prediction, and therefore suggest an adjustment to the embodiment hypothesis in which certain kinds of gestures are more effective than others.

Practical Implications

The present findings have implications for two principles of multimedia instructional design: the signaling (or cueing) principle (Mayer & Fiorella, 2014; van Gog, 2014; Xie, Mayer, Wang, & Zhou, 2019) and the embodiment principle (Mayer, 2014a; Mayer & DaPra, 2012). The *signaling principle* (also called the *cueing principle*) states that people learn better from multimedia lessons when the key material is highlighted. Forms of visual cueing include inserting arrows that point to, putting a spotlight on, or changing the color of the part of the graphic being described in the narration. The present findings help establish that another form of successful signaling (or cueing or highlighting) is for an onscreen agent to point to the specific part of graphic that she is talking about. This new addition to the collection of cueing tools is supported by the large effect size (i.e., greater than $d = 1$) created by empowering an onscreen agent with the ability to engage in specific pointing gestures. Thus, we can elaborate on the signaling to state people learn better from multimedia lessons when onscreen agents point to the relevant part of the graphic as they talk about it in a multimedia lesson.

The embodiment principle is that people learn better from multimedia lessons when an onscreen agent engages in humanlike gesturing, movement, eye contact, and facial expression. Forms of gesturing include deictic (or pointing) gestures in which the onscreen agent points to the graphic while talking about it, beat gestures in which the onscreen agent moves her hands in sync with the rhythmical pulsation of speech (Goldin-Meadow, 2003; McNeil, 2005), iconic gestures in which the onscreen agent mimics an actual movement corresponding to what is being said, and metaphorical gestures in which an onscreen agent gestures in a way that symbolizes what is being said. The present study helps distinguish between two types of deictic gesturing—SPGs, in which the onscreen agent points to the specific element in the graphic as she talks about it, and GPGs, in which she points in the general direction of graphic as she talks about parts in it. This suggests an important boundary condition for the embodiment principle in that specific pointing gestures are more effective in guiding visual attention and promoting learning than GPGs or NPGs. In short, this work helps pinpoint which kinds of humanlike gesturing are most effective for onscreen agents in multimedia instruction. In particular, instructional designers should empower pedagogical agents with pointing gestures, which is in sync with the agents' description of important elements in a graphic.

This work has implications for the role of gesture in instructional scenarios beyond pedagogical agents, such as face-to-face slideshow lectures in classrooms or instructional video presented in MOOCs, as supplements to courses, or in informal instruction found on the Internet. Thus, this work helps address the demand for effective distance learning and online courses in large undergraduate courses.

Limitations and Future Directions

Like all studies, ours has limitations that should be considered. The target sample was mainly college-age women so studies with other kinds of samples would add to the generalizability of our conclusions. Although this study involved multiple measures, including both immediate and delayed tests of retention and transfer as well eye-tracking measures that required running participants one at a time, the study is limited by investigating one short lesson in the domain of biology that is presented in a lab environment. However, the learning material of previous studies on embodiment principle mainly included the formation of lightning (e.g., Craig et al., 2015; Twyford & Craig, 2013), multistep proportional word problems (e.g., Atkinson, 2002; Lusk & Atkinson, 2007), and how solar cells work (e.g., Mayer & DaPra, 2012), which belong to different subject matter areas. Considering the embodiment principle may be influenced by the characteristic of stimuli (Baylor & Kim, 2009), in order to investigate the robustness of the findings, it would be useful to determine whether the same pattern of results would occur with different subject matter, longer lessons, and in a school setting. What's more, in this study we mainly focused on the influence of agent gestures on allocation of attention to the signaled material, retention test, and transfer test score. In the future, studies can also explore the influence of agent gestures of agent on allocation of attention to nonsignaled material and retention of nonsignaled idea, which can be helpful to better understand the selective effects of gestures on learning.

In addition, in the present study, the agent in the general pointing gesture group used eye gaze, body rotation, and general pointing gestures at the same time to guide learners' attention to the direction of illustration. However, the agents in the general pointing gesture group used in the previous studies were not the same. For example, like this study, in some studies the agent also used eye gaze, body rotation, and gestures as general pointing gestures to direct learners' attention to the illustration (e.g., Lusk & Atkinson, 2007). In other studies, the agent only used general pointing gestures, not including eye gaze and body rotation (e.g., Dunsworth & Atkinson, 2007; Yung & Paas, 2015). Thus, future study can explore whether the specific pointing gestures can still have better leaning performance, better attention distribution, and better learning perception in comparison to other kinds of general pointing gestures.

Another limitation is that the study focused on four kinds of gesturing displayed by one agent. In future research it would be useful to examine whether the superiority of specific pointing gestures (i.e., deictic gestures) remains for other kinds of onscreen agents and when compared against a natural combination of beat, iconic, and metaphorical gesturing as well as each type separately. Researchers have isolated three other types of gestures beyond pointing gestures that are often used in human communication, for example, *iconic gestures* (i.e., referring to concrete objects), *metaphoric gestures* (i.e., referring to abstract concepts), and *beat gestures* (i.e., used to show emphasis; Gawne et al., 2009). Future study can explore the role of these three other gestures in learning. It may be important to ensure that NPGs are naturally linked to the narration as they would be in normal conversation (Goldin-Meadow, 2003; McNeil, 2005). To fully understand effects of different kinds of gestures on learning, this study manipulated each gesture alone, which makes it difficult to know how the integration of different kinds of gestures affects learning. Given that teachers may produce a variety of gestures in class, it is necessary for future study to address this question.

The present study used a computer-generated character as the onscreen pedagogical agent. Because of the limitation of technology, the movements of animated agent included in this study were relatively simple and were not as real as those of a real person. Future studies should include the use of a real person as the agent to explore whether the benefits of SPGs extend to human onscreen agents in instructional videos, in which the person can exhibit more real, complex, and diverse human movements. This work can be useful given the increasing popularity of instructional videos (Fiorella & Mayer, in press).

References

Anderson, R. C., & Biddle, W. B. (1975). On asking people questions about what they are reading. *Psychology of Learning and Motivation, 9,* 89–132. http://dx.doi.org/10.1016/S0079-7421(08)60269-8

Atkinson, R. K. (2002). Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology, 94,* 416–427. http://dx.doi.org/10.1037/0022-0663.94.2.416

Baylor, A. L., & Kim, S. (2009). Designing nonverbal communication for pedagogical agents: When less is more. *Computers in Human Behavior, 25,* 450–457. http://dx.doi.org/10.1016/j.chb.2008.10.008

Baylor, A. L., & Ryu, J. (2003). The effect of image and animation in enhancing pedagogical agent persona. *Educational Computing Research, 28,* 373–395.

Beege, M., Schneider, S., Nebel, S., & Rey, G. D. (2017). Look into my eyes! Exploring the effect of addressing in educational videos. *Learning and Instruction, 49,* 113–120. http://dx.doi.org/10.1016/j.learninstruc.2017.01.004

Boucheix, J. M., Lowe, R. K., Putri, D. K., & Groff, J. (2013). Cueing animations: Dynamic signaling aids information extraction and comprehension. *Learning and Instruction, 25,* 71–84. http://dx.doi.org/10.1016/j.learninstruc.2012.11.005

Clark, R. C., & Mayer, R. E. (2016). *e-Learning and the science of instruction* (4th ed.). Hoboken, NJ: Wiley.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features and redundancy. *Journal of Educational Psychology, 94,* 428–434. http://dx.doi.org/10.1037/0022-0663.94.2.428

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

De Koning, B. B., & Tabbers, H. K. (2013). Gestures in instructional animations: A helping hand to understanding non-human movements? *Applied Cognitive Psychology, 27,* 683–689. http://dx.doi.org/10.1002/acp.2937

Dirkin, K. H., Mishra, P., & Altermatt, E. (2005). All or nothing: Levels of sociability of a pedagogical software agent and its impact on student perceptions and learning. *Journal of Educational Multimedia and Hypermedia, 14,* 113–127.

Dunsworth, Q., & Atkinson, R. K. (2007). Fostering multimedia learning of science: Exploring the role of an animated agent's image. *Computers & Education, 49,* 677–690. http://dx.doi.org/10.1016/j.compedu.2005.11.010

Fiorella, L., & Mayer, R. E. (2016). Effects of observing the instructor draw diagrams on learning from multimedia messages. *Journal of Educational Psychology, 108,* 528–546. http://dx.doi.org/10.1037/edu0000065

Fiorella, L., & Mayer, R. E. (in press). What works and doesn't work with instructional video. *Computer in Human Behavior*.

Frechette, C., & Moreno, R. (2010). The roles of animated pedagogical agents' presence and nonverbal communication in multimedia learning environments. *Journal of Media Psychology, 22,* 61–72. http://dx.doi.org/10.1027/1864-1105/a000009

Gawne, L., Kelly, B. F., & Unger, A. (2009, July). *Gesture categorisation and understanding speaker attention to gesture*. Paper presented at the 2009 Conference of the Australian Linguistic Society, Melbourne, Australia.

Glenberg, A. M., & Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse Processes, 28,* 1–26. http://dx.doi.org/10.1080/01638539909545067

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.

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

Holmqvist, K., Nystrom, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. New York, NY: Oxford University Press.

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin, 137,* 297–315. http://dx.doi.org/10.1037/a0022128

Johnson, A. M., Ozogul, G., Moreno, R., & Reisslein, M. (2013). Pedagogical agent signaling of multiple visual engineering representations: The case of the young female agent. *Journal of Engineering Education, 102,* 319–337. http://dx.doi.org/10.1002/jee.20009

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

Johnson, W. L., & Lester, J. C. (2016). Face-to-face interaction with pedagogical agents, twenty years later. *International Journal of Artificial Intelligence in Education, 26,* 25–36. http://dx.doi.org/10.1007/s40593-015-0065-9

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

Macedonia, M., Müller, K., & Friederici, A. D. (2011). The impact of iconic gestures on foreign language word learning and its neural substrate. *Human Brain Mapping, 32,* 982–998. http://dx.doi.org/10.1002/hbm.21084

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511811678

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

Mayer, R. E. (2014a). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 43–71). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139547369.005

Mayer, R. E. (2014b). Principles based on social cues in multimedia learning: Personalization, voice, image, and embodiment. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 345–370). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139547369.017

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

Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 279–315). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139547369.015

McGregor, K. K. (2008). Gesture supports children's word learning. *International Journal of Speech-Language Pathology, 10,* 112–117. http://dx.doi.org/10.1080/17549500801905622

McGregor, K. K., Rohlfing, K. J., Bean, A., & Marschner, E. (2009). Gesture as a support for word learning: The case of under. *Journal of Child Language, 36,* 807–828. http://dx.doi.org/10.1017/S0305000908009173

McNeil, D. (2005). *Gesture & thought*. Chicago, IL: University of Chicago Press. http://dx.doi.org/10.7208/chicago/9780226514642.001.0001

Moreno, R., Reislein, M., & Ozogul, G. (2010). Using virtual peers to guide visual attention during learning. *Journal of Media Psychology, 22,* 52–60. http://dx.doi.org/10.1027/1864-1105/a000008

Ozcelik, E., Arslan-Ari, I., & Cagiltay, K. (2010). Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior, 26,* 110–117. http://dx.doi.org/10.1016/j.chb.2009.09.001

Park, S. (2015). The effects of social cue principles on cognitive load, situational interest, motivation, and achievement in pedagogical agent multimedia learning. *Journal of Educational Technology & Society, 18,* 211–229.

Ryu, J., & Baylor, A. L. (2005). The psychometric structure of pedagogical agent persona. *Technology, Instruction, Cognition and Learning, 2,* 291–314.

Shiban, Y., Schelhorn, I., Jobst, V., Hörnlein, A., Puppe, F., Pauli, P., & Mühlberger, A. (2015). The appearance effect: Influences of virtual agent features on performance and motivation. *Computers in Human Behavior, 49,* 5–11. http://dx.doi.org/10.1016/j.chb.2015.01.077

Singer, M. A., & Goldin-Meadow, S. (2005). Children learn when their teacher's gestures and speech differ. *Psychological Science, 16,* 85–89. http://dx.doi.org/10.1111/j.0956-7976.2005.00786.x

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. http://dx.doi.org/10.1016/j.chb.2018.07.019

Twyford, J., & Craig, S. D. (2013). Virtual humans and gesturing during multimedia learning: An investigation of predictions from the temporal contiguity effect. In T. Bastiaens & G. Marks (Eds.), *Proceedings of E-Learn 2013–World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (pp. 2145–2149). Las Vegas, NV.

Valenzeno, L., Alibali, M. W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology, 28,* 187–204. http://dx.doi.org/10.1016/S0361-476X(02)00007-3

van Gog, T. (2014). The signaling (or cuing) principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 263–278). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139547369.014

Veletsianos, G., & Russell, G. S. (2014). Pedagogocal agents. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (4th ed., pp. 759–770). New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4614-3185-5_61

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,* 250–268. http://dx.doi.org/10.1037/edu0000221

Wang, F. X., Li, W. J., Xie, H. P., & Liu, H. S. (2017). Is pedagogical agent in multimedia learning good for learning? A meta-analysis. *Advances in Psychological Science, 25,* 12–28. http://dx.doi.org/10.3724/SP.J.1042.2017.00012

Xie, H., Mayer, R. E., Wang, F., & Zhou, Z. (2019). Coordinating visual and auditory cueing in multimedia learning. *Journal of Educational Psychology, 111,* 235–255. http://dx.doi.org/10.1037/edu0000285

Yung, H. I., & Paas, F. (2015). Effects of cueing by a pedagogical agent in an instructional animation: A cognitive load approach. *Journal of Educational Technology & Society, 18,* 153–160.

(*Appendices follow*)

Appendix A

Learning Material and 22 Idea Units in Script

Chemical synaptic transmission between neurons mainly occurs in the presynaptic membrane, synaptic gap, and the postsynaptic membrane (1). The transfer process needs more steps to complete (2).

Action potentials generated by the presynaptic neurons are transmitted to the presynaptic membrane of nerve terminals (3). The arrival of action potentials induces depolarization of presynaptic membrane (4). Thus, this intensifies the voltage gated Ca^{2+} channel on the presynaptic membrane, and permeability of Ca^{2+} is enhanced (5). At this point, Ca^{2+} in the extracellular enters into the presynaptic membrane through the channel, which leads to increasing the concentration of Ca^{2+} in the presynaptic membrane (6). The entry of Ca^{2+} may prompt the synaptic vesicle to move to the presynaptic membrane (7), and synaptic vesicle fuses with presynaptic membrane, then a cleft appears in the presynaptic membrane (8). The neurotransmitter in the synaptic vesicle is released into the synaptic gap through the role of the cell (9). These neurotransmitters arrive at the postsynaptic membrane by diffusion (10) and are combined with specific receptors on the postsynaptic membrane (11). The combination of neurotransmitters and receptors changes ion's permeability of the postsynaptic membrane, and some ion channels open (12). Ions begin to move across the membrane, for example, Na^{+} flows into the postsynaptic membrane (13), and changes the membrane potential of the postsynaptic membrane, which eventually leads to the postsynaptic potential depolarization or super polarization (14).

In order to compensate for the reduction in the number of synaptic vesicles (15), new vesicles will be reproduced under the action of the related proteins on the presynaptic membrane (16). The released neurotransmitter has an inactivation mechanism, it mainly includes three ways: First is enzyme degradation (17). Neurotransmitters that combined with receptor in the synaptic cleft, are rapidly degraded by neurotransmitter enzyme (18). Second is the diffusion (19). That is, a part of neurotransmitters leaves the synapse through passive diffusion (20). Third is to reuptake (21). That is, another part of neurotransmitters is reingested in the presynaptic membrane (22).

(*Appendices continue*)

Appendix B

Transfer Test and Answers

1. It is known that cobra venom is rich in neurotoxin. Do you know what is the poisoning mechanism of cobra bite? (The total score is two points.)

   (A) The venom can combine with specific receptors on the postsynaptic membrane, which makes neurotransmitters released by the presynaptic membrane not work for receptor (one point).

   (B) Corresponding neurotransmitter enzymes can't break down the cobra neurotoxin, which makes the toxin continue to work (one point).

2. Nerve impulses are passed from presynaptic membrane to synaptic gap to the postsynaptic membrane. Can the transfer be reversed? Why? (The total score is three points.)

   (A) It can't (one point).

   (B) Because only the presynaptic membrane has the synaptic vesicle and neurotransmitters (one point).

   (C) Only the postsynaptic membrane has the specific receptors of neurotransmitter (one point).

3. What factors can affect the process of chemical synaptic transmission? (The total score is three points.)

   (A) The amplitude and the schedule of action potential (one point).

   (V) Concentration of Ca^{2+} (one point).

   (C) The inactivation of neurotransmitter (one point).

4. According to the process of chemical synaptic transmission, if the neurotransmitter can pass nerve chemical information, then what conditions must it have? (maximum of three points).

   (A) Presynaptic neuron has the precursor and the enzyme system, which were used to synthesize neurotransmitters and then stored in vesicles (one point).

   (B) Releasing transmitters relies on presynaptic membrane depolarization and Ca^{2+} entering into the presynaptic membrane (one point).

   (C) Transmitter can be applied to specific receptors on the postsynaptic membrane and then exert its physiological effect (one point).

   (D) Synapse has the enzyme that can make neurotransmitter deactivated or has the neurotransmitter removal mechanism (one point).

   (E) Synaptic transmission function of transmitter can be strengthened or blocked by agonist or antagonist of neurotransmitter receptor (one point).

5. Please describe characteristics of receptor activity according to your understanding. (The total score is six points.)

   (A) Specificity (one point). Receptor has the identification function, namely certain receptors can only be combined with the corresponding special chemicals and generate specific physiological effects (one point).

   (B) Saturation (one point). Although it is variable, the number of receptors is always limited, which determines that its combination of a chemical molecular is also limited (one point).

   (C) Reversible (one point). Receptors can be separated from the receptor after they are combined with specific molecular chemical and play their physiological effects (one point).

Received July 9, 2018
Revision received January 22, 2019
Accepted January 29, 2019 ■