Journal of Educational Psychology
2017, Vol. 109, No. 5, 653–665

0022-0663/17/$12.00 http://dx.doi.org/10.1037/edu0000161

It's All a Matter of Perspective: Viewing First-Person Video Modeling Examples Promotes Learning of an Assembly Task

Logan Fiorella
University of Georgia

Tamara van Gog and Vincent Hoogerheide
Utrecht University and Erasmus University Rotterdam

Richard E. Mayer
University of California Santa Barbara

The present study tests whether presenting video modeling examples from the learner's (first-person) perspective promotes learning of an assembly task, compared to presenting video examples from a third-person perspective. Across 2 experiments conducted in different labs, university students viewed a video showing how to assemble an 8-component circuit on a circuit board. Students who viewed the assembly video recorded from a first-person perspective performed significantly better than those who viewed the video from a third-person perspective on accuracy in assembling the circuit in both experiments and on time to assemble the circuit in Experiment 1, but not in Experiment 2. Concerning boundary conditions, the perspective effect was stronger for more complex tasks (Experiment 1), but was not moderated by imitating the actions during learning (Experiment 1) or explaining how to build the circuit during the test (Experiment 2). This work suggests a perspective principle for instructional video in which students learn better when video reflects a first-person perspective. An explanation based on embodied theories of learning and instruction is provided.

Keywords: video, modeling examples, multimedia learning, perspective taking, embodied cognition

Consider an instructional video showing how to perform a manual task, such as how to construct a circuit on a circuit board. The main goal of this study is to examine techniques for improving the effectiveness of instructional videos, particularly the role of the perspective from which the video is recorded (i.e., first-person or third-person). In two experiments, we examine whether students learn better from an instructional video recorded from a first-person perspective, and whether there are boundary conditions for any perspective effects.

There is rapidly growing interest in the use of video modeling examples for instruction within formal (e.g., online courses) and informal (e.g., YouTube) educational settings, likely due to their convenience, relatively low cost, and high accessibility. A video modeling example involves a human model demonstrating and/or explaining to a learner how to perform a task (van Gog & Rummel, 2010). For example, a student taking an online statistics course may watch videos of an instructor solving problems on a whiteboard, or a person may watch a YouTube video of someone modeling how to tie a necktie or how to play a musical instrument. However, despite their wide implementation, there is relatively little systematic research investigating how to effectively design video lessons.

Observational Learning From Video Modeling Examples

Much of the existing research on learning from modeling examples concerns the effects of different characteristics of the human models (or animated agents) on learning. For example, in a classic study, Schunk, Hanson, and Cox (1987) manipulated the gender of the model and whether the model used an automatic mastery strategy or a more effortful coping strategy to solve math problems. More recent research has further explored the effects of model characteristics, including the model's gender (Hoogerheide, Loyens, & van Gog, 2016) and the model's age and expertise (Hoogerheide, van Wermeskerken, Loyens, & van Gog, 2016). In addition, design issues for instructional video that have been addressed recently include the visibility of the model's face (Kizilcec, Bailenson, & Gomez, 2015; van Gog, Verveer, & Verveer, 2014), the availability of gaze and gesture cues provided by the model (Ouwehand, van Gog, & Paas, 2015), the visibility of the model's hands in a motor task (Castro-Alonso, Ayres, & Paas, 2015; Marcus, Cleary, Wong, & Ayres, 2013), and whether the

This article was published Online First November 10, 2016.

Logan Fiorella, Department of Educational Psychology, University of Georgia; Tamara van Gog and Vincent Hoogerheide, Department of Education, Utrecht University, and Department of Psychology, Education, and Child Studies, Erasmus University Rotterdam; Richard E. Mayer, Department of Psychological and Brain Sciences, University of California Santa Barbara.

Tamara van Gog was supported by a Vidi grant from the Netherlands Organization for Scientific Research (NWO 452-11-06). Richard E. Mayer was supported by grants from the Office of Naval Research (N000141110225 and N0001416112046). We thank Arie Koster, Amanda Seaborne, and Kayla Shires for their assistance with Experiment 1, and Susan Ravensbergen, Kirsten Versijde, and Charly Eielts for their assistance with Experiment 2.

Correspondence concerning this article should be addressed to Logan Fiorella, Department of Educational Psychology, University of Georgia, 110 Carlton Street, Athens, GA 30602. E-mail: lfiorella@uga.edu

model physically draws out diagrams by hand during a lesson (Fiorella & Mayer, 2016). In short, past research has focused primarily on the effects of manipulating the appearance of the model and the model's actions provided in modeling examples.

The current study focuses on a largely ignored but pervasive design feature of video modeling examples: the perspective from which the video is recorded. Although perspective is typically not an issue in lecture-style modeling examples, in which the model is standing next to a screen on which slides are projected illustrating each step in the task (cf. Fiorella & Mayer, 2016; Hoogerheide, van Wermeskerken, et al., 2016; Ouwehand et al., 2015), it may play a role in demonstrations in which objects are being manipulated (cf. Castro-Alonso et al., 2015; Marcus et al., 2013; van Gog et al., 2014). Thus, perspective is a potentially important design consideration for instruction involving concrete manipulatives, commonly used to teach math and science concepts (e.g., Marley & Carbonneau, 2015).

In the current study, we tested whether students would benefit more from observing instructional videos from first-person perspective—that is, with the model performing the task from the perspective of the person observing the task—than from a third-person perspective. We also tested whether the potential effects of perspective would depend on the complexity of the to-be-learned task, and whether engaging in common and effective learning strategies—imitating in Experiment 1 and explaining in Experiment 2—would compensate for the expected detrimental effects of a third-person perspective on performance. Examining potential interactions among task complexity, learning strategies, and instructional methods is valuable because it provides insight into the robustness and generalizability of the findings. Although there is little research investigating the effects of perspective in educational videos (e.g., Lindgren, 2012), basic research in cognitive science supports the proposal that processing material from a first-person perspective may provide important cognitive benefits.

Perspective and Observational Learning

Observing the actions of others can be a powerful way to learn, likely because of the evolutionary benefits of observing and (when the outcome is desirable) imitating other people's actions (Bandura, 1977, 1986; Paas & Sweller, 2012; Sweller & Sweller, 2006). In observational learning, learners must actively interpret the actions of a human model by constructing a cognitive representation of the modeled behavior that is integrated with their prior knowledge (Bandura, 1986). Some have further proposed that this process is facilitated via activation of the mirror neuron system, which generally involves the idea that brain areas activated when performing actions are also activated when observing others perform those actions (Rizzolatti & Craighero, 2004; van Gog, Paas, Marcus, Ayres, & Sweller, 2009). When observing to-be-performed actions from the third-person perspective, learners must mentally transform the representation into their own perspective, such as by translating the model's view of left to their own left. Although humans have the unique ability to take the spatial perspective of others, such mental transformations can be cognitively demanding (Hegarty & Waller, 2004; Kessler & Thomson, 2010). This extraneous load on working memory may be reduced when the model demonstrates the task from the observer's own point of view.

Basic research in cognitive science supports a facilitative effect for processing visuomotor information from the first-person (compared to third-person) perspective. In a study by Vogt, Taylor, and Hopkins (2003), participants were asked to perform a simple hand action after being primed with pictures of hands performing either congruent or incongruent actions presented from the first- or third-person perspective. When participants were provided with a preview of the hand's start position before viewing the prime, only participants who viewed the primes from the first-person perspective were faster at performing the action when the prime displayed a congruent action compared to an incongruent action. The authors concluded that viewing body parts presented in the first-person perspective activates motor planning processes in the observer, which enhances the processing of the visual information associated with the prepared actions.

Kelly and Wheaton (2013) found further support for the notion that first-person perspective enhances motor planning and judgment. Participants were shown images of hands performing movements with tools from either a first-person or third-person perspective, and they were asked to judge the outcome of the action. Action judgments were fastest and most accurate when stimuli were viewed from the first-person perspective, again suggesting that actions are better represented when viewed from the observer's own perspective.

Next to motor planning, there is evidence from research using functional MRI (fMRI) that participants are prepared for later imitation because observing actions activates the motor neurons they would use when performing the actions (Jackson, Meltzoff, & Decety, 2006). Participants viewed video clips of simple hand and foot actions presented from the first- or third-person perspective. Some participants watched the videos passively and others imitated the actions. Behavioral data indicated that response latency to imitate the actions was shorter for the first-person perspective. Further, fMRI data indicated more activity in the left sensory-motor cortex (which would be active when executing the movement oneself) for the first-person perspective compared to third-person perspective, even when participants passively observed and did not imitate the actions. These data are consistent with embodied views of cognition (Barsalou, 2008; Wilson, 2002), which posit that human perception, cognition, and action are closely linked and grounded in one's interactions with the physical world. That is, the sensory-motor system appears more involved in processing actions from the first-person perspective, whereas the third-person perspective requires visuospatial transformations that consume limited processing capacity.

Further evidence that such transformations take time (and may result in errors) comes from a study involving visual perspective taking by Kockler and colleagues (2010). Participants were asked to make judgments about the spatial location of a static or dynamic object from their own perspective (first-person) or from the perspective of a virtual character (third-person). Results indicated that judgments were faster and more accurate when participants were asked to report the location from their own perspective. Further, fMRI data indicated that judgments of the dynamic objects from the first-person perspective resulted in increased activation in the intraparietal sulcus (IPS), an area involved in action preparation. Thus, viewing dynamic stimuli from the first-person perspective appears to improve performance by inducing a readiness to act.

Many other basic behavioral and neuroscience studies support a beneficial effect of viewing, imitating, and judging actions observed from the first-person perspective (e.g., Lorey et al., 2009; Maeda, Kleiner-Fisman, & Pascual-Leone, 2002; Surtees & Apperly, 2012; Vogeley & Fink, 2003). Similarly, the spatial cognition and navigation literatures demonstrate the high cognitive demands associated with spatial perspective taking, showing that that performance typically decreases as the angular disparity between the first-person and target viewpoint increases (e.g., Kozhevnikov, Motes, Rasch, & Blajenkova, 2006; Richardson, Montello, & Hegarty, 1999).

Unfortunately, however, research on the consequences of those findings from basic cognitive science for education and training is scarce. That is, prior research has mainly looked at effects of perspective on performance, not on learning (i.e., later performance in the absence of the observed stimuli; for an exception on spatial learning, see Richardson et al., 1999). Although prior research in educational psychology has involved video lessons presented from a first-person perspective (e.g., Ayres, Marcus, Chan, & Qian, 2009) or a third-person perspective (e.g., Arguel & Jamet, 2009), these studies did not focus on comparing the effects of video lessons presented via different perspectives. One exception is an experiment by Lindgren (2012), in which students interacted within a virtual safety training simulation from either first-person perspective or the perspective of a virtual character (i.e., third-person). Results indicated that participants who received the first-person perspective training performed better on a diagramming task, had better memory for the tasks of the simulation, committed fewer errors, and showed less help-seeking than participants who received the third-person perspective training. Lindgren concluded that virtual environments provide a unique ability to help students adopt a more embodied learning stance, allowing students to interact with learning material from their own point of view.

Overall, the available research evidence suggests a facilitation effect for processing dynamic visual information from the first-person (as opposed to third-person) perspective—consistent with the claim of embodied theories of cognition (Barsalou, 2008; Wilson, 2002) that the first-person perspective uniquely serves to shape one's cognitive representations of space and action. Accordingly, viewing materials from a third-person perspective requires learners to generate additional visuospatial transformations in order to translate observed actions into their own perspective, which creates extraneous cognitive load that consequently impairs performance. Open questions remain regarding the applicability of this basic finding to educational settings in which the focus is on learning outcomes, including potential boundary conditions associated with features of the to-be-learned task and actions of the student during learning.

The Present Study

The main aim of the current study was to investigate the hypothesis derived from the literature reviewed above, that presenting video modeling examples of an assembly task from the performer's (first-person) perspective would result in better learning (as assessed by speed and accuracy of subsequent assembly performance) than presenting videos from the third-person perspective. We conducted two experiments (in two different labs), in which university students viewed narrated video examples showing a model's hands performing an assembly task involving electric circuits. Half of the students viewed videos presented from the third-person perspective (third-person group), whereas the other half viewed videos presented from the first-person perspective (first-person group). Then, all students assembled the circuits on their own (from their perspective). According to the hypothesis, a main effect of perspective was expected in both experiments, with the first-person perspective outperforming the third-person perspective (i.e., faster and more accurate assembly).

A second aim of this study was to further explore the conditions under which video perspective influences subsequent assembly performance. Experiment 1 tested whether the effects of perspective are moderated by task complexity (within-subjects) and whether learners imitated the model during learning (between-subjects). Using only the complex tasks, Experiment 2 tested whether the effects of perspective were moderated by whether learners gave a verbal explanation while they assembled the circuit (between-subjects).

With regard to task complexity (Experiment 1), it was expected that the hypothesized beneficial effects of the first-person perspective would show primarily on complex tasks. That is, on simple tasks, which involve fewer interacting elements, overall working memory load is lower (Sweller, Ayres, & Kalyuga, 2011), and any additional processing demands imposed by the third-person perspective could be accommodated without hampering learning. The expected interaction between perspective and task complexity should also correspond to students' subjective ratings of mental effort during learning, with the highest levels of mental effort occurring when students view high-complexity tasks from the third-person perspective.

As for imitation (in Experiment 1), it was hypothesized that imitating the steps during example study might reduce reassembly time and effort, and boost test performance, for both perspectives compared to no imitation (i.e., main effect of imitation), because it would lead to deeper example processing and allow learners to practice during example study. Whereas fundamental research shows that imitation might be easier when seeing a first-person than a third-person view (Watanabe, Higuchi, & Kikuchi, 2013), it was expected that imitation might compensate for the expected negative effects of the third-person perspective (i.e., interaction effect of perspective and imitation). Although imitation is not necessary for observational learning to occur, it can aid in the process of converting symbolic codes acquired through observation into appropriate actions (Bandura, 1986). Having performed the actions (from the first-person perspective) allows for consolidating the first-person action in memory instead of the observed third-person action. Moreover, performing the actions oneself during learning may aid in transforming the observed third-person actions into first-person action representations. Without imitation, this transformation has to be made mentally. Performing the action during imitation, however, allows the learner to partially offload that mental transformation onto the external environment (e.g., by rotating the objects), thereby reducing working memory demands (Kirsh & Maglio, 1994). So, if assembly test performance is slower and less accurate in the third-person perspective condition than in the first-person perspective condition, then having made that translation during learning would boost their test performance compared to the no imitation third-person condition.

A similar expectation applied to explaining (in Experiment 2). Based on research on learning by explaining, which indicates that generating explanations is an effective learning strategy (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013; Fiorella & Mayer, 2015a, 2015b), it was expected that instructing participants that they would have to explain how to build the circuit afterward, might boost their performance compared to no explaining instruction under both perspectives (i.e., main effect of explaining). That is, knowing that they would have to give the explanation themselves later on, might result in deeper processing of the example, and especially the model's verbal explanation. Moreover, this "imitation" of the verbal explanation by the model (which was always from the first-person perspective)—by giving the same explanation to another (fictitious, nonpresent) student while assembling the circuit—might guide their assembly test performance and help compensate for detrimental effects on test performance in the third-person perspective condition (i.e., interaction effect between perspective and explaining). That is, similarly to imitating during learning, explaining during reassembly might help learners better align the model's actions and verbal instructions with their own perspective, particularly when the model's actions are presented from the third-person perspective. Taken together, the two experiments address whether engaging in learning strategies alleviates the increased processing demands expected from viewing instructional videos from the instructor's perspective.

Experiment 1

Method

Participants and design. The participants were 105 university students from the Psychology Subject Pool of a university in the United States who participated to fulfill a course requirement. The mean age of participants was 19.30 years ($SD = 1.32$), and there were 73 women and 32 men. Participants were randomly assigned to one of four conditions, based on two between-subjects factors—perspective of the instructional videos (first-person or third-person) and whether or not students imitated the video model's actions during learning (imitate or no-imitate). There were 26 students in the first-person/imitate group, 26 in the first-person/no-imitate group, 25 in the third-person/imitate group, and 28 in the third-person/no-imitate group. The groups did not significantly differ in terms of average age, number of women/men, handedness, or prior experience (as indicated by a self-report checklist described below). Task complexity (low or high) served as a within-subjects factor and was counterbalanced across conditions.

Materials. The paper-based materials consisted of a consent form, a demographics questionnaire, and a mental effort rating scale.[1] The consent form described the details of the study, informed participants that they would be videotaped during the experiment and that their privacy was protected, and included a place for them to sign. The demographics form asked participants to provide their age, gender, and handedness. Students also rated their relevant prior experience by placing a check mark next to each of eight items that apply to them, such as "I have taken a college-level course in physics," "I have worked on a circuit board," "I have installed a new light switch or electrical outlet," and "I know the difference between serial and parallel circuits."

The mental effort rating scale (Paas, 1992) asked participants to rate how much mental effort they invested while completing a particular task (e.g., watching an instructional video, building an electric circuit). Students recorded their response on a 9-point scale ranging from *Extremely low mental effort* to *Extremely high mental effort*. This common form of assessing mental effort has been shown to be sensitive enough to detect objective variations in task complexity (Ayres, 2006; Paas & Van Merrienboer, 1994).

The learning task materials consisted of a model electrical circuit kit—called Electronic Snap Circuits (by Elenco)—designed to teach students about how electrical circuits work. Students learn how to build electrical circuits by connecting (i.e., "snapping") different components (e.g., batteries, resistors, wires, LED lights) to a circuit board and to each other. In the current study, students learned how to build two circuit configurations—a low complexity circuit (shown in Figure 1) and a high complexity circuit (shown in Figure 2). As shown in the figure, both circuit configurations contain a total of eight components. However, the high complexity circuit contains more unique components (6) than the low complexity circuit (5), and the high complexity circuit contains components that must be placed in a specific orientation in order for the circuit to work. For example, in the high complexity circuit, the red LED light must point toward the green LED light, and the green LED light must point toward the battery. There are no such orientation requirements for the components in the low complexity circuit.

There were two computer-based instructional videos—a first-person version (exemplified in Figure 3) and a third-person version (exemplified in Figure 4). The instructional videos showed a male model's hands demonstrating how to build the low-complexity and high-complexity circuits while he provided narrated instructions for each of the eight steps.

The first-person perspective video showed the model's hands, as they would appear if the observer of the video were completing the task. The third-person perspective video showed the model's hands, as they would appear if someone facing the observer were completing the task. As the model placed each of the eight components on the circuit board, the oral instructions identified a component and described where it should be placed in relation to other components on the board—for example: "Place the switch below the right end of the long wire . . ." The videos were segmented to pause after the model completed a step. Students who were assigned to one of the imitating conditions would then imitate the step using their own model circuit kit before clicking to continue the video to the next step, whereas students who were assigned to one of the none-imitating conditions would simply click to continue the video to the next step. The videos were recorded simultaneously from the first-person perspective and from the third-person perspective to create identical versions from both perspectives. The low-complexity video lasted 82 seconds (excluding pauses) and contained 94 spoken words, whereas the high-complexity video lasted 90 seconds (excluding pauses) and contained 160 spoken words.

[1] We also asked participants to complete a perspective-taking test (Hegarty & Waller, 2004) upon completion of the demographics questionnaire; however, it did not significantly correlate with any of the dependent measures, and so it was not included in the analyses.

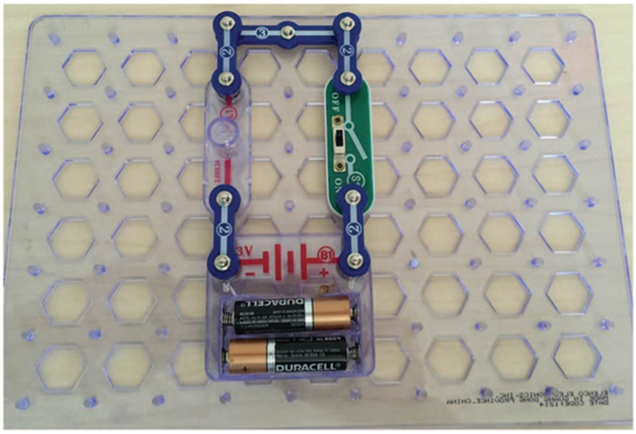

Figure 1. Low-complexity circuit. See the online article for the color version of this figure.

Figure 3. Screenshot from first-person instructional video (high-complexity task). See the online article for the color version of this figure.

Participants were assessed on their ability to assemble the low- and high-complexity circuits on their own after watching the respective instructional video. During assembly, participants were provided with the eight components needed to build the circuit along with five distractor components. Performance measures consisted of total time to assemble the circuit, accuracy at rebuilding the circuit, and frequency of three types of assembly errors. Participants were asked to assemble the circuit exactly as they saw in the video and were informed that they would be timed. Assembly time was measured from the time participants started to assemble the circuit until they stated they were finished or could not complete any more. Assembly accuracy was measured by totaling the number of correct circuit components in the correct locations and orientations on the circuit board, out of a possible 8 points for each circuit. For Experiment 1, two raters scored participants' assembly accuracy blind to experimental conditions, yielding high interrater reliability (low-complexity circuit: r = .85; high-complexity circuit: r = .82). Any discrepancies between raters were settled by consensus. For Experiment 2, there was 100% agreement between two raters based on 10% of the data, and so one rater scored the remaining data.

Assembly errors were measured by coding whether a participant committed three types of errors in their assembly of the circuit: perspective errors, location errors, or component errors. A perspective error consisted of assembling the circuit from the incorrect perspective (i.e., third-person). A location error consisted of assembling the circuit on the incorrect location on the circuit board. Finally, a component error consisted of using a component that does not make up that circuit (e.g., using a flip switch instead of a press switch). There was 100% agreement between two raters based on 10% of the data for Experiment 1, and 97% agreement between two raters based on 10% of the data for Experiment 2.

The recognition test was a lab-developed computer-based test intended to provide an additional assessment following assembly of the high-complexity circuit. The test presented participants with a series of 40 photos (one at a time) of correct or incorrect versions of the high-complexity circuit at four different orientations (i.e., first-person, 90 degrees turned left, third-person, or 90 degrees turned right). There was a delay of approximately 500 ms between each of the trials. Participants were required to determine whether each photo was the same or different from the circuit that they had just reassembled. Half of the trials were "same" trials; the other half were "different" trials, in which the photos were of a circuit

Figure 2. High complexity circuit. See the online article for the color version of this figure.

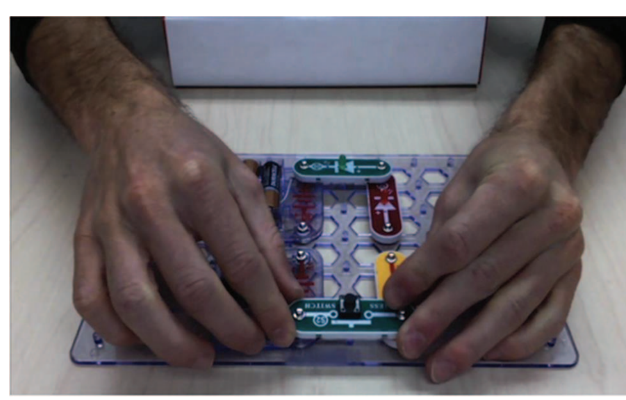

Figure 4. Screenshot from third-person instructional video (high-complexity task). See the online article for the color version of this figure.

with one component in the incorrect location (e.g., location of the resistor and red LED light switched). All trials were distributed evenly across the four orientations and presented randomly to participants using direct reaction time (RT). Performance was assessed via accuracy (out of 40) and average response time. Recognition test accuracy (but not response time) significantly correlated with assembly time, $r = -.33$, $p < .001$ and accuracy, $r = .30$, $p = .002$ for Experiment 1, providing evidence for the validity of the measure.

Apparatus. The apparatus consisted of two Dell computers with 17-in. screens, two Cyber Acoustics headphones, and two web cameras.

Procedure. The participants were randomly assigned to conditions, and were tested up to two per session in individual lab cubicles. After they provided informed consent, they completed the demographics questionnaire. Then, participants watched the first instructional video demonstrating how to build either the low- or high-complexity circuit (for counterbalancing), from either the first-person or third-person perspective (for the perspective variable). Those assigned to imitate conditions completed each step along with the video model using the electric circuit kit; those assigned to no-imitate conditions watched the video without imitating the video model (for the imitate variable). After watching the instructional video, participants completed the mental effort rating scale and then were asked to assemble the circuit on their own using the electric circuit kit. After attempting to assemble the circuit, participants again completed the mental effort rating scale. The same procedure was repeated for the second instructional video (i.e., low- or high-complexity, based on counterbalancing). The order of instructional videos was counterbalanced across conditions. After participants assembled the high-complexity circuit, participants completed the recognition test, followed by the mental effort rating scale. The total duration of the experiment was approximately 60 min.

Results and Discussion

Due to a technical issue with video recording, we do not have data for one participant's accuracy performance on the low complexity circuit. This participant's data is excluded from the relevant analyses presented below. Partial eta squared is reported as a measure of effect size, with values of .01, .06, and .14 generally representing a small, medium, and large effect size, respectively (Cohen, 1988).

Do students learn better from videos recorded from a first-person perspective than from a third-person perspective? The primary research question addressed in this study concerns whether students learn better from instructional videos presented in a first-person perspective than in a third-person perspective. Table 1 shows the mean and standard deviation for the four groups on number of correctly placed components (accuracy) on the low- and high-complexity assembly tasks. A mixed factorial analysis of variance (ANOVA) was conducted, with perspective (first-person or third-person), imitation (imitate or no-imitate), and circuit order (low-high complexity or high-low complexity) serving as between-subjects factors, and circuit complexity (low or high) serving as a within-subjects factor, and number of correctly placed components (out of 8) on the assembly tasks as the dependent measure. Consistent with predictions, there was a significant main effect of perspective, $F(1, 96) = 6.38$, $p = .013$, $\eta_p^2 = .06$, in which students correctly placed more components on the assembly tasks after viewing a first-person video ($M = 7.59$, $SD = 0.86$) than a third-person video ($M = 7.18$, $SD = 0.87$).

Table 1 also shows the mean and standard deviation for the four groups on number of seconds taken for the low- and high-complexity assembly tasks. A mixed factorial analysis of variance (ANOVA) was conducted, with perspective (first-person or third-person), imitation (imitate or no-imitate), and circuit order (low-high complexity first or high-low complexity first) serving as between-subjects factors, and circuit complexity (low or high) serving as a within-subjects factor, and total assembly time serving as the dependent measure. Consistent with predictions, there was a significant main effect of perspective, $F(1, 97) = 6.34$, $p = .013$, $\eta_p^2 = .06$, in which students completed the assembly tasks faster after viewing a first-person video ($M = 68.91$, $SD = 47.02$) than a third-person video ($M = 92.06$, $SD = 47.18$).

Table 1
Means (and SD) of Assembly Time and Accuracy per Condition for Experiment 2

	First-person perspective		Third-person perspective	
	No imitation	Imitation	No imitation	Imitation
Assembly accuracy				
Low complexity	7.73 (.67)	7.84 (.37)	7.68 (.86)	7.72 (.68)
High complexity	7.19 (1.10)	7.60 (.91)	6.57 (1.97)	6.76 (1.59)
Assembly time (s)				
Low complexity	67.38 (58.45)	45.23 (10.78)	67.61 (39.04)	54.80 (18.71)
High complexity	99.77 (99.15)	63.27 (24.57)	131.93 (95.86)	113.04 (90.81)
Recognition test				
Time	4490.66 (1,464.12)	3658.82 (1,051.98)	4195.03 (1,509.02)	4088.87 (1,109.96)
Accuracy	90.19 (14.90)	90.00 (12.37)	87.41 (15.10)	84.20 (16.55)
Mental effort (1–9)				
Example study (low complexity)	5.69 (1.78)	4.35 (1.52)	6.46 (1.35)	3.96 (1.88)
Assembly test (low complexity)	5.00 (1.92)	3.62 (1.60)	5.57 (1.67)	4.04 (2.01)
Example study (high complexity)	6.04 (1.31)	4.77 (1.82)	6.64 (1.25)	5.24 (2.26)
Assembly test (high complexity)	5.50 (1.84)	4.85 (2.22)	6.14 (1.53)	5.40 (2.16)
Recognition test	5.62 (1.42)	5.12 (1.56)	6.14 (1.51)	5.56 (1.92)

Overall, these results provide support for a perspective effect: People learn better from instructional videos recorded from a first-person perspective than from a third-person perspective. This is the primary finding of Experiment 1.

Does the perspective effect depend on the complexity of the task? A secondary question concerns whether the perspective effect favoring first-person videos is stronger for more complex assembly tasks. The ANOVA on the number of correctly placed components on the test (as summarized in Table 1) yielded a significant perspective by complexity interaction, $F(1, 96) = 4.57$, $p = .035$, $\eta_p^2 = .05$, in which the perspective effect was present in the high-complexity task (first-person video group: $M = 7.39$, $SD = 1.02$; third-person video group: $M = 6.66$, $SD = 1.79$) but not in the low-complexity task (first-person video group: $M = 7.78$, $SD = 0.54$; third-person video group: $M = 7.70$; $SD = 0.77$). Similarly, the ANOVA on the total assembly time (as summarized in Table 2) yielded a significant perspective-by-complexity interaction, $F(1, 97) = 4.79$, $p = .031$, $\eta_p^2 = .05$, in which the perspective effect was stronger for the high-complexity task (first-person video group: $M = 81.52$, $SD = 73.85$; third-person video group $M = 123.02$, $SD = 93.10$) than for the low-complexity task (first-person video group: $M = 56.31$, $SD = 43.09$; third-person group: $M = 61.57$, $SD = 31.54$). Overall, these data are consistent with the second hypothesis that the perspective effect is strong for high-complexity tasks but not for low-complexity tasks. Thus, task complexity appears to be a potential boundary condition (or moderator) for the perspective effect.

As expected, performance accuracy was significantly better on low-complexity tasks ($M = 7.74$, $SD = 0.71$) than on high-complexity tasks ($M = 7.03$, $SD = 1.43$), $F(1, 96) = 21.72$, $p < .001$, $\eta_p^2 = .19$, and assembly time was significantly shorter on low-complexity tasks ($M = 58.77$, $SD = 35.56$) than on high-complexity tasks ($M = 102.20$, $SD = 82.49$), $F(1, 97) = 27.22$, $p < .001$, $\eta_p^2 = .22$.

Does the perspective effect depend on whether students imitated the video during learning? A third question concerns whether the perspective effect favoring first-person videos is stronger when students do not have the opportunity to imitate the instructor's steps on assembly tasks. The ANOVAs showed no significant interaction between imitating and perspective for assembly accuracy $F(1, 96) < 1$, $p = .637$, or assembly time, $F(1, 97) < 1$, $p = .451$, indicating no support for the idea that imitating might compensate for the negative effects of a third-person perspective. Overall, there is no evidence for the third hypothesis that imitating during learning serves as a boundary condition (or moderator) for the perspective effect.

Imitation did, however, yield a significant main effect in which students who imitated during learning ($M = 69.30$, $SD = 47.06$) performed better on assembly time than students who not imitate ($M = 91.67$, $SD = 48.43$), $F(1, 97) = 5.93$, $p = .017$, $\eta_p^2 = .06$; however, there was not a significant main effect of imitation for assembly accuracy, $F(1, 96) = 1.30$, $p = .258$, $\eta_p^2 = .01$.

Does perspective affect the type of errors students make on assembly tasks? As a follow-up to the first research question, we also analyzed the frequency at which students made three different types of errors during assembly: perspective errors, location errors, and component errors. Perspective errors involve reassembling the circuit from the third-person perspective rather than from the first-person perspective; location errors involve reassembling the circuit on the incorrect location on the circuit board grid coordinates; and component errors involve reassembling the circuit using components that do not make up that circuit (e.g., using a flip switch instead of a press switch).

Two-sided chi-square tests were conducted to analyze the number of each type of error across perspective (first-person or third-person) and circuit complexity (low or high). For the low-complexity task, students who viewed the videos from the third-person perspective (8 out of 53, or 15.1%) were significantly more likely to commit perspective errors than students who viewed the videos from the first-person perspective (0/51, or 0%), $\chi^2(1) = 8.34$, $p = .004$. The third-person perspective group (11/53, or 20.1%) was also significantly more likely to make location errors than the first-person group, (2/51, or 3.9%), $\chi^2(1) = 6.73$, $p = .009$. The groups did not significantly differ in number of component errors (first-person: 8/51, or 15.7%; third-person: 6/53, or 11.3%; $\chi^2(1) = 0.43$, $p = .514$).

The same pattern of data was found for the high-complexity task. The third-person perspective group made significantly more perspective errors (13/53, or 24.5%) and location errors (13/53, or 24.5%) compared to the first-person group (perspective: 0/52; $\chi^2(1) = .003$, $p = .955$; location: 3/52, or 5.8%; $\chi^2(1) = 7.15$, $p = .007$), and the groups did not significantly differ on number of component errors (first-person: 13/52, or 25.0%; third-person: 13/53, or 24.5%; $\chi^2(1) = 0.003$, $p = .955$). Overall, viewing

Table 2
Means (and SD) of Assembly Time, Accuracy, Effort Ratings, and Error Types per Condition for Experiment 1

	First-person perspective		Third-person perspective	
	No explanation	Explanation	No explanation	Explanation
Assembly test				
Accuracy	6.65 (1.33)	6.60 (1.43)	5.53 (2.16)	5.38 (2.65)
Time (s)	134.45 (79.56)	145.67 (34.95)	159.63 (101.37)	172.30 (90.56)
Recognition test				
Time	4654.68 (1,930.37)	4026.33 (1,664.71)	4262.40 (1,769.96)	4003.41 (1,701.13)
Accuracy	82.94 (14.85)	80.80 (16.48)	82.17 (19.11)	79.55 (15.09)
Mental effort (1–9)				
Example study	5.77 (1.28)	6.07 (1.62)	5.83 (1.46)	6.14 (1.68)
Assembly test	5.03 (1.33)	6.20 (1.52)	5.63 (1.75)	6.00 (1.85)
Recognition test	6.06 (1.03)	5.67 (1.37)	6.37 (1.45)	5.72 (1.58)

instructional videos from the third-person perspective led to more errors related to the placement of components on the circuit board, but not more errors related to the specific components used to reassemble the circuit.

How do the treatments affect performance on the recognition test? A factorial ANOVA was conducted, with perspective (first-person or third-person), imitation (imitate or no-imitate), and circuit order (low-high complexity or high-low complexity) serving as between-subjects factors, and recognition test accuracy and response time serving as dependent measures. The analysis indicated no main effects of perspective on recognition test accuracy, $F(1, 97) = 2.25$, $p = .137$, or average response time, $F(1, 97) < 1$, $p = .771$. Further, there were no significant main effects of imitating on recognition test accuracy, $F(1, 97) < 1$, $p = .545$, or response time (although marginal), $F(1, 97) = 3.27$, $p = .074$. Finally, none of the other main effects or interactions among the factors were significant. Possibly the recognition test was not sensitive to the treatments in this study.

How do the treatments affect cognitive load? A mixed factorial ANOVA was conducted, with perspective (first-person or third-person), imitation (imitate or no-imitate), and circuit order (low-high complexity or high-low complexity) serving as between-subjects factors, circuit complexity (low or high) serving as a within-subjects factor, and self-reported cognitive load ratings as the dependent measures.

The analysis indicated no significant main effects of perspective on self-reported cognitive load throughout the experiment: after watching the low complexity video, $F(1, 97) < 1$, $p = .543$, assembling the low complexity circuit, $F(1, 97) = 1.96$, $p = .165$, watching the high complexity video, $F(1, 97) = 2.72$, $p = .102$, assembling the high complexity circuit, $F(1, 97) = 2.37$, $p = .127$, or completing the recognition test $F(1, 97) = 2.47$, $p = .119$. However, there were significant main effects of imitating, such that students who imitated along with the video model reported less cognitive load while watching the low complexity video, $F(1, 97) = 35.46$, $p < .001$, assembling the low complexity circuit, $F(1, 97) = 17.01$, $p < .001$, and watching the high complexity video, $F(1, 97) = 16.68$, $p < .001$. This difference did not reach statistical significance for assembling the high complexity circuit, $F(1, 97) = 3.29$, $p = .073$, and for completing the recognition test, $F(1, 97) = 2.84$, $p = .093$. There were no other significant main effects or interactions involving self-reported cognitive load. Overall, imitating along with the video model appears to reduce cognitive load while watching the video as well as while assembling the circuit.

Summary

Data from Experiment 1 provide initial evidence that students learn an assembly task better when instruction is presented from a first-person perspective rather than a third-person perspective. As expected, this effect was strongest for the high-complexity task, suggesting that the increased cognitive demands of the task make it more difficult for learners to overcome the detrimental effects of viewing the to-be-learned actions from the third-person perspective. Somewhat surprisingly, imitating the model's actions during learning did not appear to alleviate the influence of perspective on test performance. Experiment 2 aimed to determine whether it is possible to replicate the perspective findings and investigated whether explaining during test performance would moderate the detrimental effects of the third-person perspective video examples on test performance.

Experiment 2

The purpose of Experiment 2 was to replicate and extend the findings from Experiment 1 in another lab. First, we attempted to replicate the perspective effect using the high-complexity task from Experiment 1. Second, we tested whether a different type of learning strategy—informing participants that they would have to generate a verbal explanation during the test—might help compensate for viewing the instructional video from the third-person perspective. Since the verbal instructions provided by the model are spoken from the first-person perspective, we reasoned that informing learners that they would have to explain during reassembly might focus their attention on the model's explanation. Subsequently providing this explanation themselves might help them mentally transform actions observed from the third-person perspective into their own perspective during the test. Thus, Experiment 2 served to further test the generalizability and robustness of the perspective effect across learning contexts.

Method

Participants and design. The participants were 121 students, recruited from the subject pool of the behavioral lab of a Dutch university. One participant was excluded from the sample for failing to comply with the instructions during the experiment, leaving 120 participants. They were informed prior to signing up that the experiment would be conducted in English.[2] Participants gave informed consent during the process of signing up for the study via one of the online recruitment portals and participated either to fulfill a course requirement (psychology students, $n = 93$, 77.5%) or for a monetary reward of 5 Euro (approximately 5.43 USD at the time of writing). The mean age of participants was 21.97 years ($SD = 3.03$), and there were 68 women and 52 men. Participants were randomly assigned to one of four conditions, based on two between-subjects factors—perspective of the instructional videos (first-person or third-person), and whether or not students explained how to build the circuit to a fictitious other student during the building test (explaining or no-explaining). There were 30 students in the first-person/explaining group, 31 in the first-person/no-explaining group, 29 in the third-person/explaining group, and 30 in the third-person/no-explaining group. The groups did not significantly differ in terms of average age, proportion of men and women, or self-reported experience with circuits.

Materials. The paper-based demographics questionnaire and subjective mental effort scale were identical to those used in Experiment 1. The computer-based materials were identical to the no-imitation condition materials used in Experiment 1, with the exception that only the example video and test tasks for the high-complexity circuit were used in Experiment 2.

Apparatus. The apparatus consisted of two Hewlett-Packard computers with 22-in. screens, two Sennheiser PX30 headphones, and two web cameras.

[2] Note that this university has an international orientation; in most study programs the majority of the course literature is in English and lectures and work groups are also frequently in English.

Procedure. The participants were randomly assigned to conditions, and were tested in sessions of approximately 30 min., with a maximum of two participants per session. Participants were seated in a cubicle, which was equipped with a PC monitor on which the stimuli were presented and a webcam that was used to record their performance and explanation. First, participants provided informed consent and completed the demographics questionnaire. Participants in the no-explaining condition were then instructed that they would be watching a video example on how to build an electric circuit and that they would be asked to build it themselves afterward, whereas participants in the explain condition were instructed that they would be watching a video example on how to build an electric circuit and that they would be asked to demonstrate and explain to another student how to build it afterward. Then participants watched the instructional video demonstrating how to build the (high-complexity) circuit, from either the first-person or third-person perspective depending on assigned condition (and they all watched the video without imitating the video model). After watching the instructional video, participants rated how much effort they invested in studying it and then were asked to assemble the circuit on their own using the electric circuit kit (no-explaining condition) or to demonstrate and explain to another student how to build the circuit using the electric circuit kit (explaining condition). After attempting to assemble the circuit, participants rated how much effort they invested in this task. Finally, participants completed the recognition test, and rated how much effort they invested in this test.

Results and Discussion

Do students learn better from videos recorded from a first-person perspective than from a third-person perspective? Means and standard deviations for assembly test accuracy and assembly test time are shown in Table 2. All data were analyzed with 2×2 ANOVAs with perspective (first-person or third-person) and explanation (yes or no) as between-subjects factors, unless indicated otherwise.

In line with our main hypothesis and replicating the findings from Experiment 1, there was a significant main effect of perspective on assembly accuracy, $F(1, 116) = 10.64$, $p < .001$, $\eta_p^2 = .08$, with participants who had observed examples from the first-person perspective ($M = 6.62$, $SD = 1.37$) outperforming participants who had observed examples from the third-person perspective ($M = 5.46$, $SD = 2.39$). However, in contrast to our hypothesis and the findings from Experiment 1, the difference between the solution time of the first-person perspective group ($M = 139.97$, $SD = 61.54$) and the third-person perspective group ($M = 165.97$, $SD = 95.51$) did not reach statistical significance, $F(1, 116) = 3.12$, $p = .080$, $\eta_p^2 = .03$, in Experiment 2. Overall, there is partial support for a replication of the perspective effect found in Experiment 1.

Does the perspective effect depend on whether students explained what they were doing on the assembly test? Explanation instructions did not improve accuracy on the assembly test, $F(1, 116) < 1$, $p = .781$, and did not compensate for third-person perspective effects, as there was no significant interaction between perspective and explanation, $F(1, 116) < 1$, $p = .879$. Similarly, for assembly time, there was no main effect of explanation, $F(1, 116) < 1$, $p = .417$, nor an interaction effect between explanation and perspective, $F(1, 116) < 1$, $p = .961$. Overall, it appears that explaining during test performance was not a boundary condition (or moderator) for the perspective effect, and was not an effective technique for improving performance.

Does perspective affect the type of errors students make on assembly tasks? As in Experiment 1, we analyzed whether there were differences between the first-person and third-person perspective conditions in the number of students who made perspective errors, location errors, and component errors during the assembly test, using chi-square tests. In the third-person perspective condition, more students (34 out of 59, or 57.6%) made a perspective error than in the first-person condition (0 out of 61, or 0%), $\chi^2(1) = 49.05$, $p < .001$ (2-sided). However, there were no differences in the number of students who made location errors (in contrast to Experiment 1), $\chi^2(1) < 1$, $p = .594$ (first-person: 13/61 or 21.3%; third-person: 15/59, or 25.4%), or component errors (in line with Experiment 1), $\chi^2(1) < 1$, $p = .946$ (first-person: 41/61, or 67.2%; third-person: 40/59, or 67.8%).

How do the treatments affect recognition test performance? As in Experiment 1, there were no main effects of perspective on recognition test accuracy, $F(1, 116) < 1$, $p = .738$, or average response time on the correct trials, $F(1, 116) < 1$, $p = .552$. Further, there was no significant main effect of explaining on recognition test accuracy, $F(1, 116) < 1$, $p = .431$, or response time on the correct trials, $F(1, 116) = 1.88$, $p = .173$, and no significant interaction on recognition test accuracy, $F(1, 116) < 1$, $p = .937$, or response time on the correct trials, $F(1, 116) < 1$, $p = .569$.

How do the treatments affect cognitive load? Analysis of the self-reported mental effort invested in example study showed no main effect of perspective, $F(1, 116) < 1$, $p = .814$, or explaining, $F(1, 116) = 1.16$, $p = .283$, nor an interaction effect, $F(1, 116) < 1$, $p = .983$. The analysis of mental effort invested in the assembly test, did show a main effect of explaining, $F(1, 116) = 6.72$, $p = .011$, $\eta_p^2 = .05$, indicating—as one would expect—that participants who explained during the assembly test ($M = 6.10$, $SD = 1.68$) reported higher effort than participants who did not explain ($M = 5.33$, $SD = 1.57$). There was no main effect of perspective on effort invested in the assembly test, $F(1, 116) < 1$, $p = .499$, nor an interaction between perspective and explaining, $F(1, 116) = 1.83$, $p = .179$.

On the recognition test, explaining seemed to have a significant effect in the opposite direction: participants who had explained on the assembly test, reported lower effort investment on the recognition test ($M = 5.69$, $SD = 1.47$) than participants who had not explained ($M = 6.21$, $SD = 1.25$), $F(1, 116) = 4.33$, $p = .040$, $\eta_p^2 = .04$; however, since the test of the overall model was not significant, this effect should be interpreted with caution. There was no main effect of perspective on effort invested in the recognition test, $F(1, 116) < 1$, $p = .473$, nor an interaction between perspective and explaining, $F(1, 116) < 1$, $p = .625$. Overall, the recognition test may not be a sensitive measure.

General Discussion

Empirical Contributions

Across two experiments conducted in different labs, students learned better from instructional videos recorded from a first-person perspective than a third-person perspective as indicated by better accuracy (significant with a medium effect size in Experiment 1, $\eta_p^2 = .06$, and Experiment 2, $\eta_p^2 = .08$) and faster solution time (significant in Experiment 1, $\eta^2 = .06$, but not significant in Experiment 2, $\eta_p^2 = .02$) on an assembly test. An important boundary condition identified in Experiment 1 is that the perspective effect was strong for high-complexity assembly tasks but not for low-complexity assembly tasks. The effect of perspective on complex tasks was found whether or not students imitated the instructor during learning (in Experiment 1) and whether or not students engaged in explaining during the assembly test (in Experiment 2). Overall, this study extends basic empirical research on the facilitative role of a first-person perspective viewpoint to the design of video modeling examples of an assembly task.

Theoretical Contributions

The present study was based on predictions from embodied theories of cognition, which posit that human thought and action is deeply grounded in one's personal sensory-motor experiences of the physical world (Barsalou, 2008; Wilson, 2002). Accordingly, the first-person perspective is assumed to be critical in shaping one's internal representations of observed spatial relations and actions. We predicted that observing to-be-performed actions from the first-person perspective would facilitate the construction of a more accurate mental representation of those actions, and result in better subsequent performance, compared to observing to-be-performed actions from the third-person perspective. Findings from both experiments provided support for this prediction, demonstrating that students were generally more accurate, faster, and made fewer errors on an assembly task after viewing instructional videos presented in the first-person perspective.

Furthermore, Experiment 1 supported our second prediction that the perspective effect would be strongest for the high-complexity task compared to the low-complexity task. Observing to-be-performed actions from the third-person perspective presumably requires students to generate additional visual-spatial transformations to convert the information into their own perspective. Tasks relatively low in complexity do not excessively tax learners' limited working memory resources, leaving ample resources for making such mental transformations. However, this is different for tasks relatively high in complexity, which require students to represent a greater number of interacting elements in working memory. On such complex tasks, having to make these transformations may overload students' processing capacity and result in impaired learning.

The present study did not support predictions that imitation during example study (Experiment 1) and explaining during the building test (Experiment 2) would compensate for the negative effect of the third-person perspective on test performance. Experiment 1 indicated that imitation led to faster assembly time and reduced subjective reports of invested mental effort, but did not influence assembly accuracy. This suggests that imitation may have led to a practice effect that increased efficiency during subsequent assembly, although it did not improve test performance. This lack of effect on test performance is interesting in light of other studies that have shown that—at least in relatively short learning phases—example study followed by practice problem solving is not more effective than example study only (Leahy, Hanham, & Sweller, 2015; van Gog & Kester, 2012; van Gog et al., 2015), even when the practice opportunity is additional (i.e., not replacing an example study opportunity; Baars, van Gog, De Bruin, & Paas, 2014). This seems to underline the notion that imitation is not strictly necessary for observational learning to occur (Bandura, 1986), but note that imitation may become important in longer training sessions, to refine and automate performance (for which imitation is effective, as shown here by the reduced time on task and effort investment). More importantly in light of our present study, the experience of imitating a video model did not appear to help students overcome the detrimental effects on test performance of observing the video from the third-person perspective compared to the first-person perspective.

Similarly, in Experiment 2, knowing that one had to explain during the assembly test and actually giving an explanation to a (fictitious, nonpresent) other student, was not enough to boost performance generally (i.e., in both perspective conditions), or to counteract the cognitive demands of observing to-be-performed actions from the third-person perspective. Previous research has shown—in line with findings on self-explaining (Wylie & Chi, 2014) and peer tutoring (Roscoe & Chi, 2007)—that explaining to fictitious, nonpresent others on video camera improves learning, as evidenced by later test performance (Fiorella & Mayer, 2013, 2014; Hoogerheide, Loyens, & van Gog, 2014; Hoogerheide, Deijkers, Loyens, Heijltjes, & van Gog, 2016). In the present study, however, we investigated effects of explaining during the test itself. It is possible that the beneficial effects on knowledge restructuring that are often found to result from explaining, only manifest themselves at a later point in time.

Practical Contributions

Although the use of video modeling examples within formal and informal settings is growing rapidly, there is a paucity of rigorous empirical research to inform educators and instructional designs on how to design video lessons effectively. The present study provides preliminary evidence for a design principle of instructional videos that can be called the *perspective principle*: people learn better when instructional videos are recorded from a first-person perspective rather than a third-person perspective. This principle appears to apply most strongly for videos depicting complex tasks. Whether it applies only to learning from modeling examples on manual assembly tasks, such as assembling the components of an electric circuit, or also to other types of modeling examples or video instructions, is a question for future research to address. Our study also shows that the perspective effect is not remedied by engaging in generative learning strategies such as imitation and explaining. Although we cannot rule out the possibility that other strategies might be more successful, this strongly suggests that it is better to prevent using third-person perspective videos and create first-person videos whenever possible.

The practical relevance of our findings is strengthened by the fact that we replicated the perspective effect in two different labs (in different nations with different subject populations), indicating that the perspective effect is robust. This cross-national collaboration reflects the idea that replication is a crucial aspect of educational research (Makel & Plucker, 2014; Shavelson & Towne, 2002). Overall, the findings suggest that video modeling examples can be enhanced by presenting to-be-performed actions from the learner's own perspective.

Limitations and Future Directions

This study involved a short video on a single topic presented in a lab environment with learning assessed on an immediate test. Further work is needed to determine whether the perspective effect can be found with videos on other topics, in authentic educational settings, and on delayed tests. The perspective effect may be applicable across a wide range of domains, such as other assembly tasks, and more broadly, other types of complex motor tasks, such as when learning movements in sports, dance, or in playing a musical instrument. In academic learning, the perspective effect may extend to teaching topics in STEM domains, such as in the use of physical and virtual models to teach complex spatial relations of molecules in chemistry, or in the use of concrete manipulatives teach abstract math concepts. It may also apply to viewing other types of instructor movements, such as by viewing gestures or drawings from the first-person perspective, or to other lesson formats, such as a series of static images of an instructor manipulating objects. It might also be interesting to investigate whether the perspective effect would be stronger when viewing dynamic videos than when observing static images, as the videos might result in stronger motor-neuron activation.

The present study generally did not find effects of perspective on the recognition test or on subjective mental effort, and no complexity effect on effort in Experiment 1. The recognition test may not have been sensitive enough for detecting an influence of viewing materials from different perspectives, given that recognizing completed circuits is somewhat distinct from building a circuit on one's own, and performance was quite high in all conditions. One possibility is that the information from the instructional videos is represented in memory as actions, and therefore the recognition test did not capture learners' action-based representations.

The subjective mental effort ratings capture the overall cognitive load experienced by a learner while viewing an example or completing a task (Paas, Tuovinen, Tabbers, & Van Gerven, 2003). Given that we see all kinds of tasks being performed from a third-person perspective on a daily basis, it is not that surprising that learners do not experience higher cognitive load when studying third-person examples. On the building test, however, one might expect that having to translate from the observed third-person perspective to a first-person perspective would impose higher cognitive load and require more effort. One drawback of the fact that overall load is measured, however, is that we do not know from which cognitive processes it originates; participants in two different conditions can experience the same amount of cognitive load, while the processes from which it originates can be beneficial for learning or performance in one condition (as evidenced by higher test performance) and detrimental for learning in the other condition (as evidenced by lower test performance). We can conclude though, that the first-person perspective was more efficient, given that better test performance is reached with similar levels of effort invested in example study and test performance (van Gog & Paas, 2008). With regard to task complexity in Experiment 1, the effort ratings were somewhat higher for the more complex task, but not significantly higher. Possibly, this is due to the fact that both tasks required eight steps to be memorized. In future research, continuous and more objective measures of cognitive load, such as dual-task measures (Brünken, Plass, & Leutner, 2003) or physiological measures such as EEG (Antonenko, Paas, Grabner, & van Gog, 2010) or eye tracking (van Gog & Jarodzka, 2013) may help attain more insight into cognitive processing demands associated with viewing video examples from first- or third-person perspective.

Research should also continue to explore potential boundary conditions and moderating factors of the perspective effect. For example, individual differences such as prior knowledge, spatial ability, and working memory capacity may moderate the benefits of viewing instructional videos from the first-person perspective. That is, the perspective effect may be strongest for learners with low prior knowledge, low spatial ability, or low working memory capacity, because they do not have sufficient cognitive capacity to mentally represent and convert actions from the third-person perspective to their own perspective. The current study included a measure of spatial perspective taking; however, it was not predictive of performance on the assembly task, and so could not be explored as a potential moderator. Testing the perspective effect with more content-rich materials would also allow researchers to better explore the role of learners' prior knowledge in learning from first- and third-person perspective. In short, research is needed to clarify the generalizability of the perspective effect within different educational contexts.

Conclusion

Overall, this study contributes toward a theoretical understanding of how students learn from instructional videos and a practical understanding of how to help students learn from instructional videos. Presenting instructional videos from the learner's perspective (as opposed to from a third-person perspective) appears to better support the construction of appropriate visuospatial representations during learning, thereby resulting in better subsequent task performance. As the use of video in education continues to accelerate, this basic empirical finding offers important implications for instructional design, potentially applicable across a wide range of learning environments.

References

Antonenko, P., Paas, F., Grabner, R., & van Gog, T. (2010). Using electroencephalography (EEG) to measure cognitive load. *Educational Psychology Review, 22,* 425–438. http://dx.doi.org/10.1007/s10648-010-9130-y

Arguel, A., & Jamet, E. (2009). Using video and static pictures to improve learning of procedural contents. *Computers in Human Behavior, 25,* 354–359. http://dx.doi.org/10.1016/j.chb.2008.12.014

Ayres, P. (2006). Using subjective measures to detect variations of intrinsic cognitive load within problems. *Learning and Instruction, 16,* 389–400. http://dx.doi.org/10.1016/j.learninstruc.2006.09.001

Ayres, P., Marcus, N., Chan, C., & Qian, N. (2009). Learning hand manipulative tasks: When instructional animations are superior to equivalent static representations. *Computers in Human Behavior, 25,* 348–353. http://dx.doi.org/10.1016/j.chb.2008.12.013

Baars, M., van Gog, T., De Bruin, A., & Paas, F. (2014). Effects of problem solving after worked example study on primary school children's monitoring accuracy. *Applied Cognitive Psychology, 28,* 382–391. http://dx.doi.org/10.1002/acp.3008

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84,* 191–215.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59,* 617–645. http://dx.doi.org/10.1146/annurev.psych.59.103006.093639

Brünken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist, 38,* 53–61. http://dx.doi.org/10.1207/S15326985EP3801_7

Castro-Alonso, J. C., Ayres, P., & Paas, F. (2015). Animations showing Lego manipulative tasks: Three potential moderators of effectiveness. *Computers & Education, 85,* 1–13. http://dx.doi.org/10.1016/j.compedu.2014.12.022

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

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14,* 4–58. http://dx.doi.org/10.1177/1529100612453266

Fiorella, L., & Mayer, R. E. (2013). The relative benefits of learning by teaching and teaching expectancy. *Contemporary Educational Psychology, 38,* 281–288. http://dx.doi.org/10.1016/j.cedpsych.2013.06.001

Fiorella, L., & Mayer, R. E. (2014). Role of expectations and explanations in learning by teaching. *Contemporary Educational Psychology, 39,* 75–85. http://dx.doi.org/10.1016/j.cedpsych.2014.01.001

Fiorella, L., & Mayer, R. E. (2015a). Eight ways to promote generative learning. *Educational Psychology Review*. Advance online publication. http://dx.doi.org/10.1007/s10648-015-9348-9

Fiorella, L., & Mayer, R. E. (2015b). *Learning as a generative activity*. New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781107707085

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

Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence, 32,* 175–191. http://dx.doi.org/10.1016/j.intell.2003.12.001

Hoogerheide, V., Deijkers, L., Loyens, S. M. M., Heijltjes, A., & van Gog, T. (2016). Gaining from explaining: Learning improves from explaining to fictitious others on video, not from writing to them. *Contemporary Educational Psychology, 44/45,* 95–106. http://dx.doi.org/10.1016/j.cedpsych.2016.02.005

Hoogerheide, V., Loyens, S. M. M., & van Gog, T. (2014). Effects of creating video-based modeling examples on learning and transfer. *Learning and Instruction, 33,* 108–119. http://dx.doi.org/10.1016/j.learninstruc.2014.04.005

Hoogerheide, V., Loyens, S. M. M., & van Gog, T. (2016). Learning from video modeling examples: Does gender matter? *Instructional Science, 44,* 69–86. http://dx.doi.org/10.1007/s11251-015-9360-y

Hoogerheide, V., van Wermeskerken, M., Loyens, S. M. M., & van Gog, T. (2016). Learning from video modeling examples: Content kept equal, adults are more effective models than peers. *Learning and Instruction, 44,* 22–30. http://dx.doi.org/10.1016/j.learninstruc.2016.02.004

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

Kelly, R. L., & Wheaton, L. A. (2013). Differential mechanisms of action understanding in left and right handed subjects: The role of perspective and handedness. *Frontiers in Psychology, 4,* 957. http://dx.doi.org/10.3389/fpsyg.2013.00957

Kessler, K., & Thomson, L. A. (2010). The embodied nature of spatial perspective taking: Embodied transformation versus sensorimotor interference. *Cognition, 114,* 72–88. http://dx.doi.org/10.1016/j.cognition.2009.08.015

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic action from pragmatic action. *Cognitive Science, 18,* 513–549. http://dx.doi.org/10.1207/s15516709cog1804_1

Kizilcec, R. F., Bailenson, J. N., & Gomez, C. J. (2015). The instructor's face in video instruction: Evidence from two large-scale field studies. *Journal of Educational Psychology, 107,* 724–739. http://dx.doi.org/10.1037/edu0000013

Kockler, H., Scheef, L., Tepest, R., David, N., Bewernick, B. H., Newen, A., . . . Vogeley, K. (2010). Visuospatial perspective taking in a dynamic environment: Perceiving moving objects from a first-person-perspective induces a disposition to act. *Consciousness and Cognition, 19,* 690–701. http://dx.doi.org/10.1016/j.concog.2010.03.003

Kozhevnikov, M., Motes, M. A., Rasch, B., & Blajenkova, O. (2006). Perspective-taking vs. mental rotation transformations and how they predict spatial navigation performance. *Applied Cognitive Psychology, 20,* 397–417. http://dx.doi.org/10.1002/acp.1192

Leahy, W., Hanham, J., & Sweller, J. (2015). High element interactivity information during problem solving may lead to failure to obtain the testing effect. *Educational Psychology Review, 27,* 291–304. http://dx.doi.org/10.1007/s10648-015-9296-4

Lindgren, R. (2012). Generating a learning stance through perspective-taking in a virtual environment. *Computers in Human Behavior, 28,* 1130–1139. http://dx.doi.org/10.1016/j.chb.2012.01.021

Lorey, B., Bischoff, M., Pilgramm, S., Stark, R., Munzert, J., & Zentgraf, K. (2009). The embodied nature of motor imagery: The influence of posture and perspective. *Experimental Brain Research, 194,* 233–243. http://dx.doi.org/10.1007/s00221-008-1693-1

Maeda, F., Kleiner-Fisman, G., & Pascual-Leone, A. (2002). Motor facilitation while observing hand actions: Specificity of the effect and role of observer's orientation. *Journal of Neurophysiology, 87,* 1329–1335.

Makel, M. C., & Plucker, J. A. (2014). Facts are more important than novelty: Replication in the education sciences. *Educational Researcher, 43,* 304–316. http://dx.doi.org/10.3102/0013189X14545513

Marcus, N., Cleary, B., Wong, A., & Ayres, P. (2013). Should hand actions be observed when learning hand motor skills from instructional animations? *Computers in Human Behavior, 29,* 2172–2178. http://dx.doi.org/10.1016/j.chb.2013.04.035

Marley, S. C., & Carbonneau, K. J. (2015). How psychological research with instructional manipulatives can inform classroom learning. *Scholarship of Teaching and Learning in Psychology, 1,* 412–424. http://dx.doi.org/10.1037/stl0000047

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

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive load approach. *Journal of Educational Psychology, 84,* 429–434. http://dx.doi.org/10.1037/0022-0663.84.4.429

Paas, F., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. *Educational Psychology Review, 24,* 27–45. http://dx.doi.org/10.1007/s10648-011-9179-2

Paas, F., Tuovinen, J., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38,* 63–71. http://dx.doi.org/10.1207/S15326985EP3801_8

Paas, F., & Van Merrienboer, J. J. G. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review, 6,* 351–371. http://dx.doi.org/10.1007/BF02213420

Richardson, A. E., Montello, D. R., & Hegarty, M. (1999). Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Memory & Cognition, 27,* 741–750. http://dx.doi.org/10.3758/BF03211566

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27,* 169–192. http://dx.doi.org/10.1146/annurev.neuro.27.070203.144230

Roscoe, R. D., & Chi, M. T. H. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. *Review of Educational Research, 77,* 534–574. http://dx.doi.org/10.3102/0034654307309920

Schunk, D. H., Hanson, A. R., & Cox, P. D. (1987). Peer-model attributes and children's achievement behaviors. *Journal of Educational Psychology, 79,* 54–61. http://dx.doi.org/10.1037/0022-0663.79.1.54

Shavelson, R. J., & Towne, L. (Eds.). (2002). *Scientific research in education*. Washington, DC: National Academy Press.

Surtees, A. D., & Apperly, I. A. (2012). Egocentrism and automatic perspective taking in children and adults. *Child Development, 83,* 452–460.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4419-8126-4

Sweller, J., & Sweller, S. (2006). Natural information processing systems. *Evolutionary Psychology, 4,* 434–458. http://dx.doi.org/10.1177/147470490600400135

van Gog, T., & Jarodzka, H. (2013). Eye tracking as a tool to study and enhance cognitive and metacognitive processes in computer-based learning environments. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 143–156). New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4419-5546-3_10

van Gog, T., & Kester, L. (2012). A test of the testing effect: Acquiring problem-solving skills from worked examples. *Cognitive Science, 36,* 1532–1541. http://dx.doi.org/10.1111/cogs.12002

van Gog, T., Kester, L., Dirkx, K., Hoogerheide, V., Boerboom, J., & Verkoeijen, P. P. J. L. (2015). Testing after worked example study does not enhance delayed problem-solving performance compared to restudy. *Educational Psychology Review, 27,* 265–289. http://dx.doi.org/10.1007/s10648-015-9297-3

van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist, 43,* 16–26. http://dx.doi.org/10.1080/00461520701756248

van Gog, T., Paas, F., Marcus, N., Ayres, P., & Sweller, J. (2009). The mirror neuron system and observational learning: Implications for the effectiveness of dynamic visualizations. *Educational Psychology Review, 21,* 21–30. http://dx.doi.org/10.1007/s10648-008-9094-3

van Gog, T., & Rummel, N. (2010). Example-based learning: Integrating cognitive and social-cognitive research perspectives. *Educational Psychology Review, 22,* 155–174. http://dx.doi.org/10.1007/s10648-010-9134-7

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

Vogeley, K., & Fink, G. R. (2003). Neural correlates of the first-person-perspective. *Trends in Cognitive Sciences, 7,* 38–42. http://dx.doi.org/10.1016/S1364-6613(02)00003-7

Vogt, S., Taylor, P., & Hopkins, B. (2003). Visuomotor priming by pictures of hand postures: Perspective matters. *Neuropsychologia, 41,* 941–951. http://dx.doi.org/10.1016/S0028-3932(02)00319-6

Watanabe, R., Higuchi, T., & Kikuchi, Y. (2013). Imitation behavior is sensitive to visual perspective of the model: An fMRI study. *Experimental Brain Research, 228,* 161–171. http://dx.doi.org/10.1007/s00221-013-3548-7

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9,* 625–636. http://dx.doi.org/10.3758/BF03196322

Wylie, R., & Chi, M. T. H. (2014). The self-explanation principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 413–432). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139547369.021

Received May 17, 2016
Revision received September 14, 2016
Accepted September 15, 2016 ■