Received: 5 March 2022 | Revised: 18 August 2022 | Accepted: 25 August 2022

DOI: 10.1002/acp.3994

SHORT PAPER

WILEY

Learning by explaining after pauses in video lectures: Are provided visuals a scaffold or a crutch?

Logan Fiorella

Department of Educational Psychology, University of Georgia, Athens, Georgia, USA

Correspondence
Logan Fiorella, Department of Educational Psychology, University of Georgia, Athens, GA, USA.
Email: lfiorella@uga.edu

Funding information
National Academy of Education/Spencer Postdoctoral Fellowship; National Science Foundation, Grant/Award Number: 1955349

Abstract

Explaining after pauses in a video lecture can be an effective learning activity, yet students need support to generate comprehensive explanations. This study tested whether providing students access to the visualizations from the video enhances explanation comprehensiveness and transfer performance. Undergraduates ($n = 112$) watched a 5-part lesson on the human kidney consisting of explanations and drawings. After each part, students typed an explanation without access to information from the video (explain group) or with access to the visualizations from the video (explain-visuals group), or they studied the visualizations and transcript from the video (restudy group). Students who explained significantly outperformed the restudy group on the transfer test ($d = .47$). The explain-visuals group generated significantly more comprehensive explanations during learning than the explain group ($d = .54$), but this did not result in significantly better performance on the subsequent transfer test ($d = .09$), suggesting the visuals may have served as a crutch more than a scaffold.

KEYWORDS

generative learning, learning by explaining, learning from video, learning from visualizations, multimedia learning

1 | INTRODUCTION

The use of video in education continues to accelerate in the wake of the COVID-19 pandemic (Adedoyin & Soykan, 2020; Ali, 2020). In recent years, researchers have produced a small but growing body of literature on research-based principles for designing video lessons (de Koning et al., 2018; Fiorella & Mayer, 2018; Mayer et al., 2020). Many of these principles focus on how the information is presented to students, such as breaking down the video into smaller parts (Spanjers et al., 2012), providing subtitles (Lee & Mayer, 2018), or using cues to direct students' attention to critical elements of the learning material (Pilegard & Fiorella, 2021). These methods can be effective for managing students' limited cognitive capacity, but they do not guarantee students will spontaneously use their available resources effectively to support learning (de Koning et al., 2011). For example, inserting pauses in between segments of a video lesson is generally more effective than showing the entire video continuously (i.e., the segmenting principle; see Mayer & Fiorella, 2022), but students may additionally benefit from explicit prompts to make sense of the learning material during those pauses (Fiorella, 2022). The present study explores the effects of prompting learners to generate *explanations* after each segment of a complex video lecture about the human kidney.

2 | EXPLAINING AS A GENERATIVE ACTIVITY

According to generative learning theory (Fiorella & Mayer, 2015, 2016), explaining is a *generative activity* because it encourages students to select essential information from a lesson, organize it into a coherent structure, and integrate it with their existing knowledge. For example, when learning about the kidney, students may *select* key components of how blood is filtered in the glomerulus and *organize* them into a coherent causal sequence: blood enters the afferent arteriole, some blood is filtered into Bowman's capsule, and the rest exits the glomerulus via the efferent arteriole. Students also may *integrate* aspects of the filtration process with their existing knowledge, such as by inferring that because the glomerulus is fenestrated with small holes, only the small components of the blood pass through to Bowman's capsule.

There is substantial evidence that prompting students to generate explanations during learning results in deeper understanding than not prompting students to explain (Chi, 2022). A recent meta-analysis by Bisra et al. (2018) found a medium average effect size of $g = .55$ across 69 comparisons in favor of prompts to elicit self-explanations. The benefits of explaining hold across a range of learning contexts, including learning from text, visualizations, and worked examples. In recent years, studies have also found preliminary support for prompting students to explain after pauses in video lessons. For example, Fiorella et al. (2020) found students who wrote explanations after each part of a 5-part video lesson on the human kidney exhibited higher transfer test performance than students who either rewatched each part of the video or students who created drawings after each part of the video. Similarly, a series of experiments by Lawson and Mayer (2021) found students who wrote explanations during pauses in a 4-part animated lesson about greenhouse gasses significantly outperformed unprompted students on delayed posttests.

Despite these benefits, many students struggle to generate comprehensive explanations that fully capture the key ideas and conceptual relationships from the learning material (Chi et al., 1989; Renkl, 1997). This is important because the ability to generate high-quality explanations is a strong predictor of performance on subsequent measures of meaningful learning, such as comprehension or transfer (Chi et al., 1989; Fiorella et al., 2020). Thus, a major goal of research on learning by explaining is to identify effective and feasible ways to boost explanation quality and, consequently, learning outcomes. Prior attempts to improve explanation quality involve extensive training or providing adaptive prompting and feedback tailored to each student's responses (e.g., Berthold et al., 2009; McNamara, 2017). The present study tests a simpler method: using the provided visuals from a video lesson as a scaffold for supporting learner-generated explanations.

3 | VISUALIZATIONS AS SCAFFOLDS

Recent accounts of generative learning theory (Fiorella & Mayer, 2022)—based on related frameworks of learning from visualizations (Mayer, 2022; McCrudden & Rapp, 2017)—emphasize the complementary relationship between verbal and visual representations. In particular, visualizations support learning in large part because they scaffold generative *explanation* processes, which together result in the construction of a coherent and more accurate mental model (Butcher, 2006; Cox, 1999; Eitel et al., 2013). For example, in a study by Butcher (2006), students were asked to generate self-explanations while learning from a lesson on the human circulatory system that contained either text and provided diagrams or text alone. Students who received diagrams generated more accurate inferences in their explanations and constructed a more accurate mental model of the system than students who only studied the text. Other studies involving either provided or learner-generated visuals report a similar pattern of results (Ainsworth & Loizou, 2003; Cromley et al., 2010; Fiorella & Kuhlmann, 2020). A study by Eitel et al. (2013) provides related evidence for the visual-as-scaffold hypothesis. Students who were only briefly shown a diagram of a pulley system (for 600 ms or 2 s) exhibited better comprehension of a subsequent text on the functions of the pulley system than students who studied the text without exposure to the diagram. Together, these studies suggest visualizations are a promising candidate for improving the effectiveness of learning by explaining.

However, one potential downside of explaining external visualizations is that learners may engage in less effortful retrieval processes than if they were to explain without the visuals present (Pyc & Rawson, 2009). The retrieval practice literature indicates that students generally benefit from actively retrieving information from long-term memory more than receiving additional exposure to the material (Karpicke, 2012). However, research on learning by explaining in open- vs. closed-book formats has produced somewhat mixed results (Hiller et al., 2020; Roelle & Berthold, 2017; Waldeyer et al., 2020), suggesting it is not always clear when students should explain as a retrieval practice on their own or when students might benefit from access to (parts of) the learning material as possible retrieval supports.

4 | THE PRESENT STUDY

In the present study, college students watched a 5-part video lecture about the human kidney. The video contained a series of oral explanations and dynamic drawings created by the instructor. After each part of the video, students were prompted to type explanations of the concepts from the lesson without access to the content of the video (explain group), write explanations of the concepts from the lesson with access to the visuals presented in the video (explain-visuals group), or study the provided visuals and the transcript from that part of the video (restudy group). Then all students completed retention and transfer tests of the learning material.

In line with generative learning theory and related empirical research, I expected students prompted to explain (explain-visuals and explain groups) would outperform the restudy group on the transfer test (*learning by explaining hypothesis*). I did not have strong predictions regarding retention test performance because prior research indicates the benefits of explaining are most prominent on transfer tests, which better reflect the quality of students' mental models (Chi & Wylie, 2014).

The primary goal of the study was to test whether provided visualizations would further enhance the effectiveness of learning by

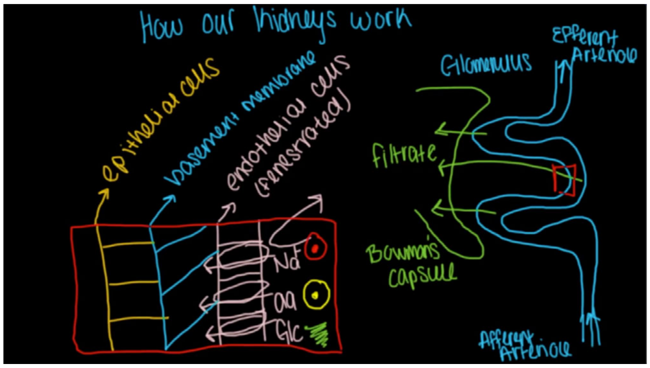

FIGURE 1 Screenshot from part 4 of video lesson on human kidney

explaining. The *visuals-as-scaffold hypothesis* posits that the explain-visuals group will generate more comprehensive explanations and exhibit higher transfer test performance than the explain group. The rationale for this prediction is that explaining provided visuals supports the construction of a more accurate mental model of the learning material. For example, when learning about the filtration of blood in the glomerulus, a provided visual (see Figure 1) can display each the vessels that enter and exit the glomerulus, and each layer of Bowman's capsule that the blood must pass through during filtration. Rather than being forced to internally visualize these relationships, students can use the provided visual as a framework from which to construct their own coherent explanation of how blood is filtered, including what is filtered and the rate of filtration. Students can then apply their constructed mental model of the filtration process to solve transfer problems, such as determining what would happen to the filtration rate if the vessels increased or decreased in diameter.

However, an alternative hypothesis is that visuals may serve as a *crutch* rather than a scaffold. One recent study in a different learning context found that provided visuals supported more comprehensive explanations initially but this did not correspond to better subsequent learning outcomes (Fiorella et al., 2021). It is possible that students may rely too heavily on the provided *external* visualization when constructing their explanation and thereby do not engage in sufficiently effortful retrieval processes to form an adequate *internal* mental model of the learning material. Thus, the *visuals-as-crutch* posits the explain-visuals group should generate more comprehensive explanations than the explain group, but that this benefit will disappear on the transfer test, when students do not have access to the provided visuals.

The visuals-as-scaffold and visuals-as-crutch hypotheses also yield different predictions regarding the relationship between the comprehensiveness of students' explanations and transfer test performance. The visual-as-scaffold hypothesis posits explanation comprehensiveness should be positively associated with transfer for both the explain-only group and explain-visuals group. In contrast, the visuals-as-crutch hypothesis posits the relationship between explanation comprehensiveness and transfer test performance will be strongest for explain-only condition and weaker (or nonexistent) for the explain-visuals condition.

Finally, as supplementary data, I used a self-report survey to explore whether the three groups differed in their perceptions of the learning material, including perceived difficulty, enjoyment, interest, and motivation to learn, as prior research indicates visualizations can provoke positive emotions and reduce perceived difficulty (Moreno, 2006).

5 | METHOD

5.1 | Participants and design

Participants were 151 undergraduates recruited from the Educational Psychology Subject Pool at a large university in the southeast United States, who participated online via Qualtrics to receive course credit. Thirty-nine participants were removed for not completing the entire study or not following instructions (e.g., watching the videos more than once; copying and pasting responses to questions from the Internet). The remaining 112 participants were assigned randomly to one of three groups: explain ($n = 38$), explain-visuals ($n = 35$), or restudy ($n = 39$). A power analysis using G*Power with the same parameters as Lawson and Mayer (2021; power $= .80$, $f = .325$, $\alpha = .05$) indicated that a total sample size of 96 was sufficient for this study. The mean age of participants was 20.1 years ($SD = 1.7$), and there were 19 men and 93 women. The groups did not differ significantly in mean age, $F(2, 111) = 0.24$, $p = .786$, proportion of men and women, $\chi^2(2) = .762$, $p = .683$, or performance on a prior knowledge test (described below), $F(2, 111) = 1.50$, $p = .227$.

5.2 | Materials

Materials consisted of a participant survey, a prior knowledge test, a video lesson on the human kidney, explanation prompts, retention and transfer tests, and a learning experience survey. All materials were adapted from Fiorella et al. (2020).

The participant survey asked students to provide their age, gender, major, and year in college. The prior knowledge test consisted of five short-answer questions ($\alpha = .63$) targeting background knowledge related to the human kidney: "How many kidneys do humans have?", "What are the main functions of the kidney?", "What is a ureter?", "What is a glomerulus?" and "What is the functional unit of the kidney?" Participants were awarded one point for each correct response. All questions required specific responses, so one rater scored all participants' responses using an established rubric.

The video lesson explained the structure and function of the kidney and was segmented into five parts. Each part consisted of oral explanations from a female instructor and supporting diagrams that were labeled and dynamically drawn by the instructor. The instructor was not visible on the screen but used a tablet to create the diagrams. The videos were adapted by Fiorella et al. (2020) from existing videos developed by the Khan Academy, a popular online platform consisting of thousands of short videos on a wide range of topics for students of all ages. The popularity of the Khan Academy (over 70 million users;

Khan Academy, 2018) reflects the increased popularity among students and teachers of using relatively short videos to supplement classroom instruction.

The total length of the lesson was 14 minutes and 32 s; it consisted of 1272 words and six unique diagrams. (Some diagrams from earlier parts of the lesson appeared again on a subsequent part of the lesson.) The five parts of the video included: (1) overview of the structure and functions of the kidney (2 min and 46 s, 252 words, two diagrams); (2) renal physiology and anatomy (3 min, 290 words, one unique diagram); (3) glomerular structure in the nephron (1 min and 45 s, 130 words, one unique diagram), (4) glomerular filtration in the nephron (2 min and 43 s, 207 words, one unique diagram) and (5) the processes of passive transport and countercurrent multiplication (4 min and 18 s, 393 words, one unique diagram). Figure 1 presents a screenshot from Part 4 of the lesson. The diagram on the right remained on the screen from Part 3 (glomerular structure), and the diagram on the left was new and unique to the content of Part 4 (glomerular filtration).

The explanation prompts were presented to the explain group and the explain-visuals group after each part of the video. Both groups responded to the same five prompts, corresponding to the five parts of the video. For example, after Part 4, participants were asked to: “Explain how blood is filtered through the vessels of the glomerulus.” The explain group responded to the prompts without access to information from the video. The explain-visuals group was shown a screenshot of the visualizations from the lesson while responding to each prompt, and each explanation prompt included the phrase “Use the picture above to…” at the beginning. For example, Figure 1 was shown to the explain-visuals group after Part 4 of the video, with the prompt, “Use the picture above to explain how blood is filtered through the vessels of the glomerulus.”

Students' explanations were scored based on their comprehensiveness by providing one point for each correct component, or idea unit, included in their explanation. Because comprehensiveness captures the number of accurate ideas and relationships included in students' explanations, it serves as a proximate indicator of generative processing (Fiorella & Mayer, 2015). For example, idea units from Part 4 of the lesson included: “vessels in the glomerulus are lined with endothelial cells,” “holes in the vessels allow smaller components to pass through,” and “large components cannot be filtered and move through the efferent arteriole.” In total, there were 51 possible idea units students could include in their explanations across the five parts of the lesson (Part 1: 8 idea units, Part 2: 9, Part 3: 9, Part 4: 7, and Part 5: 18).

Two raters independently scored the explanations for 12 participants using a shared rubric. Inter-rater reliability was high (ICC $= .94$), and so one rater scored the remaining responses.

The post-tests consisted of a retention test and a transfer test. The retention test consisted of 5 short-answer questions ($\alpha = .42$) assessing students' memory for specific information from the lesson: for example, “What is the specific function of the glomerulus?” and “What is the name of the vessel that carries blood away from the glomerulus?” All questions required specific responses, so one rater scored all participants' responses using a scoring rubric. Because the retention test consisted of a low number of items, I also used the average inter-item correlation to assess reliability (Clark & Watson, 1995). The average inter-item correlation was .14, which is just outside Clark and Watson's (1995) recommended range of .15–.50.

The transfer test consisted of 5 free-response questions ($\alpha = .46$) assessing students' ability to go beyond the lesson to generate new inferences and apply their knowledge of the kidney to new situations: (1) “Explain what happens to large components of the blood in the glomerulus,” (2) “Explain why the kidney does not remove all of the waste from the blood,” (3) “How would you adapt the nephron to support an animal that lives in very dry climates,” (4) “Explain what would happen to the rate of blood filtration in the glomerulus if the diameter of the afferent arteriole increased and the diameter of the efferent arteriole decreased,” and (5) “Explain what would happen if a patient received a drug that decreased salt permeability of the cells that line the ascending limb of the Loop of Henle.” The average inter-item correlation for the transfer test items was .15, which is within Clark and Watson's (1995) recommended range of .15–.50. Also note that it is common for transfer tests that cover a broad range of knowledge to have relatively lower internal consistency (e.g., Wiley et al., 2005). The transfer test was scored by awarding one point for each correct solution. For example, acceptable responses for Question 3 include: (a) passive transport could be more efficient in the Loop of Henle in order to improve water reabsorption; (b) more ions could be reabsorbed from the ascending limb in order to reabsorb more water by countercurrent multiplication in the descending limb; and (c) the Loop of Henle could be longer, in order to collect more water. In total, the transfer test was worth 21 possible points, though students were not expected to provide all possible solutions to each question. Two raters scored responses on the transfer test for 12 participants. Interrater reliability was high (ICC $= .96$), and so one rater scored all remaining responses.

The learning experience survey consisted of seven Likert-scale items assessing students' perceptions of the learning material on a 5-point scale from (1) strongly disagree to (5) strongly agree: “I felt that the subject matter was difficult,” “I enjoyed learning this way,” “I would like to learn this way in the future,” “I feel like I have a good understanding of the material,” “After this lesson, I would be interested in learning more about the material,” “I found the lesson to be useful to me,” and “I felt motivated to try to understand the material.”

5.3 | Procedure

I originally planned to collect data in a laboratory setting. However, due to COVID-19 restrictions, participants completed the study online via Qualtrics. They were assigned randomly to one of the three experimental groups: explain, explain-visuals, or restudy. After providing informed consent, students completed the participant survey and the prior knowledge test. Then they read instructions informing them that they would watch a 5-part video lesson on the human kidney and complete a brief learning activity after each part of the video. The instructions also informed them that they would be tested on their understanding of how the human kidney works.

Next, participants watched the video lesson and completed their respective learning activity. The explain group responded to an explanation prompt after each part of the video by typing their explanations into a text box. The explain-visuals group responded to the same prompts but were additionally provided with a screenshot of the visualizations from that part of the video to help them write their explanations. The restudy group received the same screenshots as the explain-visuals group and were additionally provided with the complete transcript from that part of the video. The transcript was presented directly below the screenshots. Participants were instructed to study the provided picture and read the transcript from the lesson before moving on to the next part of the lesson. Participants were not able to advance to the next part of the video until they spent a minimum of 30 s explaining or restudying. This minimum was set to deter students from simply clicking to the next screen without completing the explanation prompts or studying the provided information. Total time spent during the pauses in the video were recorded. Students were not able to go back to earlier parts of the video.

After completing the video and their respective learning activities, all participants completed the retention test followed by the transfer test. Participants completed the tests at their own pace by typing their responses into a text box. Finally, they completed the learning experience survey.

6 | RESULTS

The data were analyzed to test for the effects of generating explanations on learning, the role of provided visuals as a scaffold or a crutch, the role of time spent during pauses in the video, and students' perceptions of the learning materials.

6.1 | Does generating explanations enhance learning from video lectures?

Table 1 presents the means and standard deviations of the comprehensiveness of students' explanations during learning and their subsequent retention and transfer test scores. To test the learning-by-explaining hypothesis, I conducted a one-way analysis of variance (ANOVA) with planned contrast tests (explain group and explain-visuals group > restudy group), using retention and transfer test scores as dependent measures. Consistent with my hypothesis, students who generated explanations significantly outperformed the restudy group on the transfer test, $t(109) = 2.45$, $p = .016$, $d = .47$, but not on the retention test, $t(109) = 0.38$, $p = .708$, $d = .07$.

6.2 | Do provided visuals serve as a scaffold or a crutch?

To test the visuals-as-scaffold and visuals-as-crutch hypotheses, I conducted independent samples t-tests comparing the explain group and explain-visuals group on explanation comprehensiveness during learning, as well as their subsequent learning outcomes. Results indicated that the explain-visuals group generated significantly more comprehensive explanations during learning, $t(71) = 2.28$, $p = .026$, $d = .54$, but did not score significantly higher on the subsequent retention test, $t(71) = 0.25$, $p = .807$, $d = .06$, or the transfer test, $t(71) = 0.40$, $p = .690$, $d = .09$. This pattern of results favors the visuals-as-crutch hypothesis.

Next, I derived partial correlations between explanation comprehensiveness and learning outcomes, controlling for performance on the prior knowledge test. For the explain-only group, explanation comprehensiveness significantly predicted retention and transfer performance, with Pearson's coefficients of $r = .47$ ($p = .004$) and $r = .47$ ($p = .003$), respectively. For the explain-visuals group, the relationships were not statistically significant, with coefficients of $r = .16$ ($p = .362$) for retention and $r = .29$ ($p = .098$) for transfer. Overall, these findings provide additional support for the visuals-as-crutch hypothesis.

6.3 | Did the groups differ in how much time they spent during pauses?

One possibility is that differences in explanation comprehensiveness or transfer are explained by the amount of time students spent during the pauses in the video. A one-way ANOVA revealed that the three groups significantly differed in average pause time, $F(2,109) = 22.01$, $p < .001$. Tukey post-hoc tests indicated that the restudy group ($M = 62.18$ s; $SD = 36.7$) spent significantly less time than the explain group ($M = 99.43$, SD $= 43.1$, $p = .001$) and the explain-visuals group ($M = 130.5$, $SD = 52.3$, $p < .001$). Furthermore, the explain group spent significantly less time than the explain-visuals group ($p = .010$). I then examined correlations among pause times and transfer test performance for each group. Spending more time during the pauses was not significantly associated with better transfer performance for any of the groups, and was weakest for the restudy group: restudy ($r = .17$, $p = .312$), explain ($r = .27$, $p = .106$), or explain-visuals ($r = .32$, $p = .06$). Furthermore, pause times were only correlated with explanation comprehensiveness for the explain group ($r = .45$, $p = .004$) but not for the explain-visuals group ($r = .26$, $p = .135$). Taken together, this pattern provides additional support for the

TABLE 1 Means and standard deviations for learning outcomes and explanation comprehensiveness across groups

	Explanation comprehensiveness		Retention test		Transfer test	
Group	*M*	*SD*	*M*	*SD*	*M*	*SD*
Restudy	–	–	3.5	1.0	2.4	1.7
Explain	11.6	6.2	3.6	1.5	3.4	1.8
Explain-Visuals	14.9	6.0	3.6	1.2	3.2	1.8

visuals-as-crutch hypothesis: Despite spending longer during pauses and generating more comprehensive explanations, the explain-visuals group did not exhibit superior transfer performance compared to the explain group.

6.4 | Did the groups differ in their perceptions of the learning materials?

Finally, I explored whether the groups differed on their perceptions of the learning materials. Because the data did not meet assumptions of normality, I conducted nonparametric independent-samples Kruskal–Wallis tests. Results indicated the groups significantly differed in their self-reported interest in learning more about the material, $H(2) = 10.68$, $p = .005$, and their perceived usefulness of the lesson, $H(2) = 9.37$, $p = .009$. Post-hoc tests indicated that the explain-visuals group reported greater interest in learning more about the learning material than the explain group ($p = .008$) and perceived the learning material as more useful than the explain group ($p = .012$) and the restudy group ($p = .050$). Surprisingly, the restudy group also reported greater interest in learning more about the learning material than the explain group ($p = .027$). There were no other significant differences across groups on items from the learning experience survey.

7 | DISCUSSION

This study replicated and extended prior research on the benefits of writing explanations during pauses in a video lesson (Fiorella et al., 2020; Lawson & Mayer, 2021). According to generative learning theory (Fiorella & Mayer, 2016, 2022), prompting students to explain encourages them to select important information from the lesson, organize it into a coherent mental representation in working memory, and integrate it with one's existing knowledge from long-term memory. If students are able to generate comprehensive explanations that incorporate the critical elements and relationships from the lesson, they should exhibit superior transfer test performance compared to students not prompted to explain. Results from the present study were in line with basic predictions from generative learning theory. First, students who wrote explanations significantly outperformed the restudy group on the subsequent transfer test. Second, the comprehensiveness of students' explanations (for the explain-only group) was significantly positively associated with performance on the retention and transfer test.

The novel contribution of this study involved testing whether providing access to the static visuals from the video (a) enhances explanation comprehensiveness during learning, and (b) results in better transfer test performance when the visuals are no longer available. According to the *visuals-as-scaffold hypothesis*, provided visuals serve as an external scaffold, which facilitates construction of a higher-quality *internal* representation (or mental model) of the learning material. Alternatively, the *visuals-as-crutch hypothesis* posits that provided visuals may allow students to explain more effectively when the visuals are accessible, but it may not help students to better *internalize* the learning material into a coherent mental model for later use. The results of the present study favored the visuals-as-crutch hypothesis: (a) visuals facilitated explaining during learning but did not result in better subsequent transfer test performance, and (b) explanation comprehensiveness was significantly associated with transfer performance for the explain group but not for the explain-visuals group.

It is important to note that prior research showing facilitative effects of visuals on learner-generated explanations compared students who learn from text and visuals to students who learn from *only* text (Ainsworth & Loizou, 2003; Butcher, 2006; Eitel et al., 2013). The present study was different because it took place in the context of learning from video, in which all students watched the same video lesson containing oral explanations and provided visualizations from the instructor. The learner-generated explanations occurred after each part of the lesson and thus served as a *retrieval* activity (Karpicke, 2012). Provided visuals were intended to facilitate retrieval by providing students with explicit cues regarding the structure of the learning material (Eitel et al., 2013). The rationale was that generative retrieval activities depend on the extent to which students are *successful* at retrieving the relevant information to construct a quality explanation (reflected by explanation comprehensiveness). It is possible, however, that the visuals provided supportive cues at the expense of retrieval *effort* (Pyc & Rawson, 2009), in which students in the explain-visuals group were not required to invest sufficient effort retrieving the material from memory. Related research comparing explaining in open- and closed book format also highlights the challenges of balancing when and how students should be required to retrieve the learning material (Hiller et al., 2020; Roelle & Berthold, 2017), with one recent study suggesting a combination of open- and closed-book formats may be most effective (Waldeyer et al., 2020). Within the present context, a future study could explore generative activities that impose sufficient retrieval effort but also provide support to ensure quality explanations. For example, students may be asked to generate and explain their own visualizations of the learning material and subsequently receive the provided visualizations and/or explanations as feedback. Research on learning by drawing indicates that this approach may be superior to relying only on provided visuals (Van Meter, 2001; Zhang & Fiorella, 2019).

From a practical standpoint, the present research contributes to a growing literature exploring research-based principles for designing effective video lessons. The results were consistent with the generative activity principle (Mayer et al., 2020): prompting students to generate explanations after each part of an instructional video enhances learning. Adding supports like provided visuals may help students generate more comprehensive explanations in the moment but this approach may not benefit subsequent transfer performance. Interestingly, the explain-visuals group reported greater enjoyment and interest in learning more about the learning material; these potential motivational benefits should be explored in future studies that track learning and engagement over longer time periods. Nonetheless, instructors may consider other methods such as providing focused prompts or hints to guide students' explanations or prompting students to generate their own visualizations.

7.1 | Limitations and future directions

Due to COVID-19, data collection for this study was conducted online; thus, there was limited control of participants' learning environment. I took steps to screen for participants who did not follow instructions, but other factors may have influenced the quality of the data. For example, performance on the transfer test was relatively low in this study. The study is also somewhat limited because it only assessed immediate learning outcomes, though prior work has demonstrated the benefits of explaining for long-term learning (e.g., Lawson & Mayer, 2021).

Another limitation of the study is that the outcome measures only assessed learning verbally. It is possible that the benefits of explaining provided visuals would be more pronounced if the assessments required students to reason from provided visuals or draw elements of the kidney themselves. However, it should be noted that the transfer questions, though presented verbally, are intended to reflect the quality of students' mental models, which are assumed to include visuospatial elements. In other words, explaining provided visuals was expected to improve the quality of one's mental model, which should yield better transfer performance. Indeed, performance on drawing tests tends to be correlated with performance on verbal transfer tests (Fiorella & Kuhlmann, 2020). That said, explicitly prompting students to generate visualizations would provide more direct insight into the quality of students' knowledge than relying on verbal responses alone.

The retention test used in this study was also limited because it consisted of only 5 short-answer items. Although the retention test was intended to assess students' recall of some of the key ideas from the lesson, it is not as comprehensive as alternative measures, such as free recall tests. As I discuss above, the focus of this study was on transfer test performance because strategies like explaining tend to most effective for assessments of understanding rather than mere recall (Chi & Wylie, 2014). Nonetheless, including a broader measure of retention would provide a more complete picture of how explaining affects different types of learning outcomes.

Another potential criticism of the study is that it lacked additional comparison groups, such as a restudy group that only studied the visuals (without the provided transcripts), or an explanation group that explained the provided transcripts from the lesson (without the provided visuals). My primary focus was on determining whether explaining provided visuals was more effective than explaining without visuals (explain-only group vs. explain-visuals group), and I included the restudy group as a control to confirm whether explaining was more effective than an opportunity to restudy all the learning materials (transcripts and visuals). It would be valuable for future research to include additional comparison groups to clarify that (a) explaining visuals is more effective than only studying the provided visuals, and (b) that explaining visuals is more effective than explaining the provided transcripts.

Finally, future research should consider incorporating additional measures such as students' spatial ability, metacognitive judgments of learning, or cognitive load. Such measures might help explain why the benefits of explaining provided visuals did not carry over to the learning outcome measures. For example, provided visualizations may differentially serve as a scaffold or a crutch for students with different levels of spatial ability (Höffler, 2010). Furthermore, it is possible that the availability of provided visuals while explaining inflates students' judgment of their understanding because the explaining task is perceived as easier or requires less mental effort (Wiley, 2019). Consequently, students in the explain-visuals group may not have invested as much effort toward (a) internalizing the provided visuals into a coherent mental model and/or (b) attending to subsequent parts of the video. Including direct measures of students' cognitive load and metacognitive judgments in future work would help clarify these issues.

ACKNOWLEDGMENTS

I thank Robert Hebert for his assistance with data coding.

FUNDING STATEMENT

This research was supported by a grant from the National Science Foundation (1955349) and a National Academy of Education/Spencer Postdoctoral Fellowship.

CONFLICT OF INTEREST

The author has no conflicts of interest to disclose.

DATA AVAILABILITY STATEMENT

Data will be made available upon request to the author.

ORCID

Logan Fiorella https://orcid.org/0000-0002-1672-325X

REFERENCES

Adedoyin, O. B., & Soykan, E. (2020). COVID-19 pandemic and online learning: The challenges and opportunities. *Interactive Learning Environments*, 1–13. [Advance online publication].

Ainsworth, S., & Loizou, A. T. (2003). The effects of self-explaining when learning with text or diagrams. *Cognitive Science*, *27*, 669–681.

Ali, W. (2020). Online and remote learning in higher education institutes. A necessity in light of COVID-19 pandemic. *Higher Education Studies*, *10*(3), 16–26.

Berthold, K., Eysink, T. H., & Renkl, A. (2009). Assisting self-explanation prompts are more effective than open prompts when learning from multiple representations. *Instructional Science*, *37*(4), 345–363.

Bisra, K., Liu, Q., Nesbit, J. C., Salimi, F., & Winne, P. H. (2018). Inducing self-explanation: A meta-analysis. *Educational Psychology Review*, *30*, 703–725.

Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology*, *98*, 182–197.

Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*(2), 145–182.

Chi, M. T. H. (2022). The self-explanation principle in multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 381–393). Cambridge University Press.

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*, 219–243.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*, 309–319.

Cox, R. (1999). Representation construction, externalized cognition and individual differences. *Learning and Instruction*, *9*, 343–363.
Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Cognitive activities in complex science text and diagrams. *Contemporary Educational Psychology*, *35*(1), 59–74.
de Koning, B. B., Hoogerheide, V., & Boucheix, J. M. (2018). Developments and trends in learning with instructional video. *Computers in Human Behavior*, *89*, 395-398.
de Koning, B. B., Tabbers, H. K., Rikers, R. M., & Paas, F. (2011). Improved effectiveness of cueing by self-explanations when learning from a complex animation. *Applied Cognitive Psychology*, *25*(2), 183–194.
Eitel, A., Scheiter, K., Schuler, A., Nystrom, M., & Holmqvist, K. (2013). How a picture facilitates the process of learning from text: Evidence for scaffolding. *Learning and Instruction*, *28*, 48–63.
Fiorella, L. (2022). Multimedia learning from instructional video. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed.). Cambridge University Press.
Fiorella, L., & Kuhlmann, S. (2020). Creating drawings enhances learning by teaching. *Journal of Educational Psychology*, *112*(4), 811–822.
Fiorella, L., & Mayer, R. E. (2015). *Learning as a generative activity: Eight learning strategies that promote understanding*. Cambridge University Press.
Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review*, *28*(4), 717–741.
Fiorella, L., & Mayer, R. E. (2018). What works and doesn't work with instructional video. *Computers in Human Behavior*, *89*, 465–470.
Fiorella, L., & Mayer, R. E. (2022). The generative activity principle in multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed.). Cambridge University Press.
Fiorella, L., Pyres, M., & Hebert, R. (2021). Explaining and drawing activities for learning from multimedia: The role of sequencing and scaffolding. *Applied Cognitive Psychology*, *35*(6), 1574–1584.
Fiorella, L., Stull, A. T., Kuhlmann, S., & Mayer, R. E. (2020). Fostering generative learning from video lessons: Benefits of instructor-generated drawings and learner-generated explanations. *Journal of Educational Psychology*, *112*(5), 895–906.
Hiller, S., Rumann, S., Berthold, K., & Roelle, J. (2020). Example-based learning: Should learners receive closed-book or open-book self-explanation prompts? *Instructional Science*, *48*, 623–649.
Höffler, T. N. (2010). Spatial ability: Its influence on learning with visualizations—A meta-analytic review. *Educational Psychology Review*, *22*(3), 245–269.
Karpicke, J. (2012). Retrieval-based learning: Active retrieval promotes meaningful learning. *Current Directions in Psychological Science*, *21*(3), 157–163.
Khan Academy. (2018). 2018 in review: Annual report. *Khan Academy*. https://blog.khanacademy.org/2018-in-review-annual-report/
Lawson, A. P., & Mayer, R. E. (2021). Benefits of writing an explanation during pauses in multimedia lessons. *Educational Psychology Review*, *33*, 1–27.
Lee, H., & Mayer, R. E. (2018). Fostering learning from instructional video in a second language. *Applied Cognitive Psychology*, *32*(5), 648–654.
Mayer, R. E. (2022). Cognitive theory of multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 57–72). Cambridge University Press.
Mayer, R. E., & Fiorella, L. (2022). Principles for managing essential processing in multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 243–260). Cambridge University Press.
Mayer, R. E., Fiorella, L., & Stull, A. T. (2020). Five ways to increase the effectiveness of instructional video. *Educational Technology Research & Development*, *68*, 837–852.
McCrudden, M. T., & Rapp, D. N. (2017). How visual displays affect cognitive processing. *Educational Psychology Review*, *29*(3), 623–639.
McNamara, D. S. (2017). Self-explanation and reading strategy training (SERT) improves low-knowledge students' science course performance. *Discourse Processes*, *54*(7), 479–492.
Moreno, R. (2006). Does the modality principle hold for different media? A test of the method-affects-learning hypothesis. *Journal of Computer Assisted Learning*, *22*, 149–158.
Pilegard, C., & Fiorella, L. (2021). Using gestures to signal lesson structure and foster meaningful learning. *Applied Cognitive Psychology*, *35*(5), 1362–1369.
Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language*, *60*(4), 437–447.
Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science*, *21*(1), 1–29.
Roelle, J., & Berthold, K. (2017). Effects on incorporating retrieval into learning tasks: The complexity of the tasks matters. *Learning and Instruction*, *49*, 142–156.
Spanjers, I. A., Van Gog, T., Wouters, P., & Van Merriënboer, J. J. (2012). Explaining the segmentation effect in learning from animations: The role of pausing and temporal cueing. *Computers & Education*, *59*(2), 274–280.
Van Meter, P. (2001). Drawing construction as a strategy for learning from text. *Journal of Educational Psychology*, *93*, 129–140.
Waldeyer, J., Heitmann, S., Moning, J., & Roelle, J. (2020). Can generative learning tasks be optimized by incorporation of retrieval practice? *Journal of Applied Research in Memory and Cognition*, *9*, 355–369.
Wiley, J. (2019). Picture this! Effects of photographs, diagrams, animations, and sketching on learning and beliefs about learning from a geoscience text. *Applied Cognitive Psychology*, *33*(1), 9–19.
Wiley, J., Griffin, T. D., & Thiede, K. W. (2005). Putting the comprehension in metacomprehension. *Journal of General Psychology*, *132*, 408–428.
Zhang, Q., & Fiorella, L. (2019). Role of generated and provided visuals in supporting learning from scientific text. *Contemporary Educational Psychology*, *59*, 101808.

How to cite this article: Fiorella, L. (2022). Learning by explaining after pauses in video lectures: Are provided visuals a scaffold or a crutch? *Applied Cognitive Psychology*, *36*(5), 1142–1149. https://doi.org/10.1002/acp.3994