Received: 22 February 2021 | Revised: 9 July 2021 | Accepted: 2 September 2021
DOI: 10.1002/acp.3871

SHORT PAPER

WILEY

Explaining and drawing activities for learning from multimedia: The role of sequencing and scaffolding

Logan Fiorella | **Macy Pyres** | **Robert Hebert**

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

Correspondence
Logan Fiorella, Department of Educational Psychology, University of Georgia, Athens, Georgia, 30605.
Email: lfiorella@uga.edu

Funding information
National Academy of Education/Spencer Postdoctoral Fellowship; National Science Foundation, Grant/Award Number: 1955348

Summary

This study tested how different implementations of explaining and drawing activities affect learning from a multimedia science lesson. After studying a multimedia slideshow about the human respiratory system, college students ($n = 145$) were assigned randomly to one of four learning activity conditions: write explanations before drawing pictures (explain-then-draw group), draw pictures before writing explanations (draw-then-explain group), study provided explanations before drawing pictures (provided explanation-then-draw), or study provided pictures before writing explanations (provided drawing-then-explain). One week following the learning activity, all students completed post-tests of their understanding. Results from the learning activity supported the scaffolding hypothesis: students generated better explanations when they used provided (rather than their own) drawings, and they generated better quality drawings when they used provided (rather than their own) explanations. However, this difference in learning activity performance did not correspond to higher performance on the delayed post-tests. We discuss implications for how to best sequence and scaffold generative activities.

KEYWORDS
drawing, explaining, generative learning, learning strategies, multimedia

1 | INTRODUCTION

Many students struggle to understand complex multimedia science lessons. One common remedy is to improve how the learning material is *presented* to students, such as by cueing important information, breaking down the lesson into smaller parts, or placing words next to corresponding graphics (see Mayer, 2020). Yet even well-designed multimedia lessons do not guarantee students will actively *make sense* of what they are learning (e.g., de Koning et al., 2011). Thus, another approach is to prompt students to engage in *generative learning activities* (Fiorella & Mayer, 2016)–activities that encourage students to construct meaning from the lesson by organizing and integrating it with their existing knowledge.

Prior research has identified two types of generative activities shown to improve student understanding in science: *Explaining* and *drawing*. Explaining involves generating *verbal* statements (orally or in writing) to clarify the meaning of what one is learning (Wylie & Chi, 2014), such as when reading a text, studying examples, or analyzing diagrams. Explaining supports the construction of a coherent mental model by encouraging students to generate inferences, monitor their understanding, and repair gaps in their knowledge (Chi, 2000). A recent meta-analysis by Bisra et al. (2018) found an average effect size of $g = .55$, across 69 comparisons, in support of prompting students to self-explain. One potential boundary condition is that students may need guidance to generate high-quality explanations (e.g., Berthold et al., 2009).

Drawing involves creating a *pictorial* representation that depicts the critical structures and processes from the learning material, often-scientific text (Van Meter & Garner, 2005). For example, students may create drawings that depict water molecules (Leopold & Leutner, 2012), the mechanics of a tire pump (Bobek & Tversky, 2016), or the human circulatory system (Zhang & Fiorella, 2019). According to the cognitive model of drawing construction (Van Meter & Firetto, 2013), drawing supports learning because it forces students to integrate verbal and pictorial representations. For example, when reading the sentence "the pulmonary trunk splits into the right and left pulmonary arteries," students must use their existing knowledge to translate the provided text into a corresponding image depicting the structure of the pulmonary trunk and arteries. There is growing empirical support for the benefits of learning by drawing (Ainsworth & Scheiter, 2021). A recent meta-analysis by Fiorella and Zhang (2018) found drawing is generally more effective than alternative strategies like writing summaries; yet like explaining, students often need guidance to benefit from drawing activities (e.g., Van Meter, 2001).

1.1 | Integrating explaining and drawing activities

One persistent limitation of prior research is that it often compares individual generative activities to relatively weak control group activities, such as students who only read the lesson (Chi & Wylie, 2014; Fiorella & Zhang, 2018). Indeed, explaining and drawing have largely been examined separately in prior research; few studies have compared them directly (e.g., Scheiter et al., 2017), and even fewer have tested how they might be combined to further support learning (e.g., Fiorella & Kuhlmann, 2020). This is surprising considering the well-documented benefits of combining *instructor-provided* verbal and visual representations in the multimedia learning literature (Mayer, 2020). Visuals, such as diagrams or illustrations, complement verbal explanations by conveying the learning material in a more organized, efficient, and meaningful way (Stenning & Oberlander, 1995). In fact, one primary reason instructor-provided visuals are effective is because they encourage students to generate higher-quality *self-explanations* (i.e., generate more inferences) than when they learn from text alone (Ainsworth & Loizou, 2003; Butcher, 2006; Cromley et al., 2010). In short, provided visuals appear to serve as scaffold for sense-making (Eitel et al., 2013), facilitating learner-generated explanations and resulting in better understanding.

A recent study by Fiorella and Kuhlmann (2020) provides preliminary insight into the relationship between learner-generated drawings and learner-generated explanations. College students studied a lesson on the human respiratory system and then were asked to teach what they learned on video as if they were instructing a peer. Students taught by either explaining aloud (explain-only group), creating drawings on paper (draw-only group), or creating drawings on paper while simultaneously explaining aloud (draw-and-explain group). Results indicated the draw-and-explain group outperformed both the explain-only and draw-only groups (who did not differ from each other) on post-tests administered 1 week later. Critically, analyses of students' explanations indicated the draw-and-explain group generated higher quality explanations than the explain-only group, which partially explained their higher post-test performance. In other words, creating drawings facilitated explaining, which resulted in better understanding–consistent with earlier predictions from Cox (1999) on the role of generated external representations in problem solving. In contrast, the draw-and-explain group did not produce better quality drawings than the draw-only group, suggesting that explaining did not facilitate drawing.

Taken together, drawing (or viewing provided visuals) appears to support explaining, but generating an explanation may not have the same facilitative effect on drawing (or viewing provided visuals). One implication of this view is that there may be unique benefits of using visual representations to generate explanations as opposed to using explanations to generate visual representations–what we refer to as the *sequence hypothesis*.[1] The study by Fiorella and Kuhlmann (2020) was not able to test this hypothesis directly because students created drawings and explained aloud simultaneously. A related issue is whether students should receive scaffolding to support the translation from visual to verbal representations (or vice versa). Given the struggles students experience to generate quality explanations and drawings on their own (e.g., Berthold, Eysink, & Renkl., 2009; Van Meter, 2001), students may benefit from using provided explanations or visuals (rather than their own) as a scaffold for this translation process–what we refer to as the *scaffolding hypothesis*.

1.2 | The present study

The goal of this research was to explore how the sequencing and scaffolding of generative activities affects learning. In the present study, college students studied a multimedia slideshow presentation about the human respiratory system and then completed a learning activity that involved generating explanations and/or drawings. Students were assigned randomly to write explanations before drawing pictures (explain-then-draw group), draw pictures before writing explanations (draw-then-explain group), study provided explanations before drawing pictures (provided explanation-then-draw group), or study provided drawings before writing explanations (provided drawing-then-explain group). One week following the learning activity, all students completed post-tests, which included explanation, drawing, and transfer tests. We investigated how sequencing and scaffolding uniquely affects (a) explanation and drawing quality during the learning activity and (b) subsequent explanation, drawing, and transfer post-test performance. We describe our major hypotheses below.

Our two *sequence* hypotheses are based on prior research suggesting unique facilitative effects of instructor-provided visualizations on students' verbal explanations and understanding (e.g., Ainsworth & Loizou, 2003; Butcher, 2006; Cox, 1999; Fiorella & Kuhlmann, 2020):

- Sequence hypothesis 1: Students who draw pictures and then write explanations will generate higher quality explanations during the learning activity, resulting in higher subsequent post-test performance than students who write explanations and then draw pictures (draw-then-explain group > explain-then-draw group). Thus, we expect explanation quality to mediate the relationship between learning activity and post-test performance. We also expect the facilitative effect on explanation quality will be stronger for the explanation post-test than the drawing and transfer post-tests.

- Sequence hypothesis 2: Students who write explanations after studying provided visuals will exhibit higher subsequent post-test performance than students who draw pictures after studying provided explanations (provided drawing-then-explain group > provided explanation-then-draw group). We also expect the benefits of writing explanations from provided visuals will be stronger for the explanation post-test than the drawing and transfer post-tests.

Our two *scaffolding* hypotheses are based on prior research suggesting students need support to generate high-quality explanations and drawings (e.g., Berthold, Eysink, & Renkl, 2009; Van Meter, 2001):

- Scaffolding hypothesis 1: Students who write explanations after studying provided visuals will generate higher quality explanations during the learning activity, resulting in higher subsequent post-test performance than students who write explanations after creating their own drawings (provided drawing-then-explain group > draw-then-explain group). Thus, we expect explanation quality to mediate the relationship between learning activity and post-test performance. We also expect the facilitative effect on explanation quality will be stronger for the explanation post-test than the drawing and transfer post-tests.
- Scaffolding hypothesis 2: Students who draw pictures after studying provided explanations will generate higher quality drawings during the learning activity, which will exhibit higher subsequent post-test performance than students who draw pictures after writing their own explanations (provided explanation-then-draw group > explain-then-draw group). Thus, we expect drawing quality to mediate the relationship between learning activity and post-test performance. We also expect the facilitative effect on drawing quality will be stronger for the drawing post-test than the explanation and transfer post-tests.

In addition to the learning activity and post-test measures, we also asked students to complete surveys about their learning experience throughout the study, including their ratings of invested effort, perceived difficulty, helpfulness, and their level of understanding. The learning experience surveys were used as exploratory data to provide additional insight into the effects of sequencing and scaffolding generative activities.

2 | METHOD

2.1 | Participants and design

Participants were 145 undergraduates recruited from the Educational Psychology Subject Pool at a large university in the southeast United States, who participated to receive course credit. Participants were assigned randomly to one of four conditions: draw-then-explain ($n = 46$), explain-then-draw ($n = 34$), provided explanation-then-draw ($n = 32$), or provided drawing-then-explain ($n = 33$). The mean age of participants was 19.8 ($SD = 1.8$), and there were 14 men, 130 women, and 1 student who responded "other." The groups did not differ significantly in mean age, $F(3,141) = 0.45$, $p = .716$, or proportion of men and women, $X^2(6) = 6.55$, $p = .365$.

2.2 | Materials and measures

Materials consisted of a lesson on the human respiratory system, post-tests of learning outcomes, learning activity materials, a learning experience survey, a participant survey, and a spatial ability test.

The lesson on the human respiratory system (adapted from Fiorella & Kuhlmann, 2020) contained 759 words and consisted of an introductory paragraph, followed by nine main paragraphs supplemented by corresponding diagrams. The post-tests included an explanation test, a drawing test, and a transfer test. The explanation test consisted of one item asking students to explain how the human respiratory system works. This item was scored by awarding students one point for each correct component of their explanation corresponding to the inhalation, exchange, and exhalation processes. For example, correct responses to the explanation test include: When inhaling, the ribs expand outward; when at rest, the diaphragm is pointed up; and capillaries facilitate exchange of O_2 and CO_2. The drawing test consisted of two items ($\alpha = .45$) asking students to draw and label structures of the human respiratory system: (1) "Draw a picture of the respiratory system when a person inhales and label the different parts"; and (2) "Draw a picture of the exchange system and label the different parts. Because the drawing test consisted of a low number of items, we also used the inter-item correlation to assess reliability (Clark & Watson, 1995). The inter-item correlation of the two drawings items was .32, which is within Clark and Watson's (1995) recommended range of .15–.50 and suggests good internal reliability. The drawing items were scored by awarding students one point for each correctly depicted structure, such as the bronchial tubes, the thoracic cavity, or the diaphragm moving down during inhalation.

The transfer test consisted of five items ($\alpha = .44$) asking students to apply their knowledge of the human respiratory system to new situations: For example, "Suppose there is oxygen in the lungs, but the cells in the body do not get enough oxygen to make energy. What could have caused this problem?" "How could you redesign the human respiratory system to make it more effective in climates where less oxygen is in the air (such as at the top of tall mountains)?" Because the transfer test consisted of a low number of items, we again used inter-item correlations to assess the reliability of the transfer test. The average inter-item correlation among the five items was .14, which is just outside Clark and Watson's (1995) recommended range of .15–.50. The transfer items were scored by awarding students one point for each correct solution. For example, correct responses to the first transfer question include: Alveoli not sending oxygen to the capillaries; or too much carbon dioxide in the lungs. Two raters scored responses to each of the post-tests for

16 participants. Interrater reliability was good (ICC's > .77), and so the one rater scored all remaining responses.

The materials for the learning activity varied depending on experimental condition. Students in the *explain-then-draw* group were prompted to write an explanation of the inhalation, exchange, and exhalation process, and then to use their written explanation to draw pictures that depict the inhalation, exchange, and exhalation processes. Students in the *draw-then-explain* group received the same prompts but in the reverse order–draw pictures and then use their pictures to write an explanation of the inhalation, exchange, and exhalation processes. Students in the *provided explanation-then-draw* group received text-based explanations of the inhalation, exchange, and exhalation processes (excerpted directly from the lesson) and were prompted to use the provided text to draw pictures that depict each process. Students in the *provided drawing then-explain* group received pictures of the inhalation, exchange, and exhalation processes (excerpted directly from the lesson) and were prompted to use the provided pictures to write an explanation of each process. All paper-based materials were presented on 8.5 × 11 sheets of paper. We scored the accuracy of student's drawings and written explanations for the learning activity following the same scoring procedure as the explanation test and the drawing test described above. Figure 1 shows the pictures provided to the explain-from-pictures group, and Figure 2 shows an example of a student's drawings from the draw-from-explanations group.

The learning experience survey was administered three times throughout the experiment: after studying the lesson, after completing the learning activity, and after completing the post-test. The surveys contained four items asking students to rate their invested mental effort (while studying, completing the learning activity, or completing the post-test), the perceived difficulty (of the lesson, learning activity, or post-test], the perceived helpfulness [of the lesson or learning activity), and their level of understanding on a 7-point Likert-like scale. For example, the survey following the learning activity directed participants to: (1) "Please rate the amount of mental effort you invested while performing the learning activity" (very low to very high); (2) "Please rate how difficult you found the learning activity" (very easy to very hard); (3) Please rate how helpful the learning activity was for understanding how the human respiratory system works" (very unhelpful to very helpful); and (4) "Please rate how well you understand how the respiratory system works" (very poorly to very well).

The participant survey directed participants to provide their age, gender, major, and whether or not they have taken an AP or college-level biology course. They were also asked to rate their confidence and interest in learning about science by rating their level of agreement (on a 7-point scale) with the following statements: "When I take a science course (e.g., biology, physics, chemistry), I am confident I will perform well," and "I like to learn about science in my free time (e.g., books, museums, documentaries)." Finally, to assess self-reported prior knowledge, students were asked to indicate which of six statements applied to them: For example, "I know the differences among arteries, veins, and capillaries," and "I know the function of the diaphragm. Self-reported prior knowledge scores could range from 0 to 6 points.

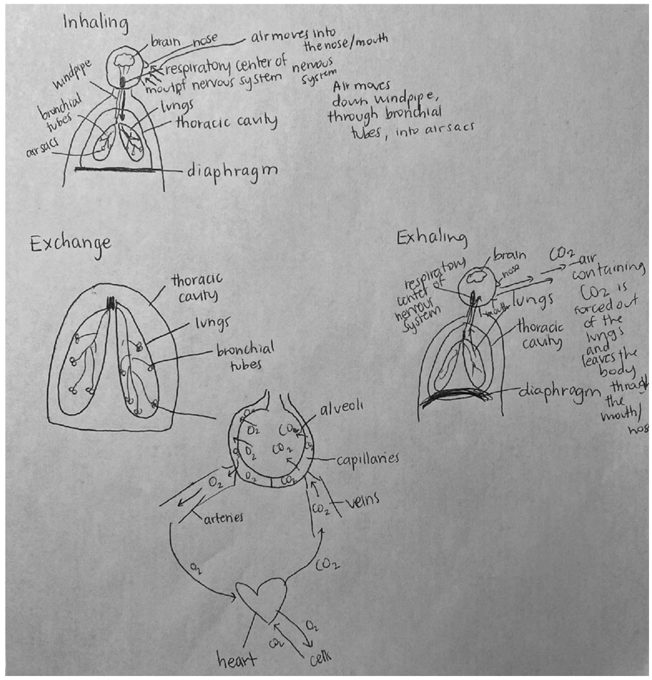

FIGURE 2 Example of a student's drawings during the learning activity (provided drawing-then-explain group)

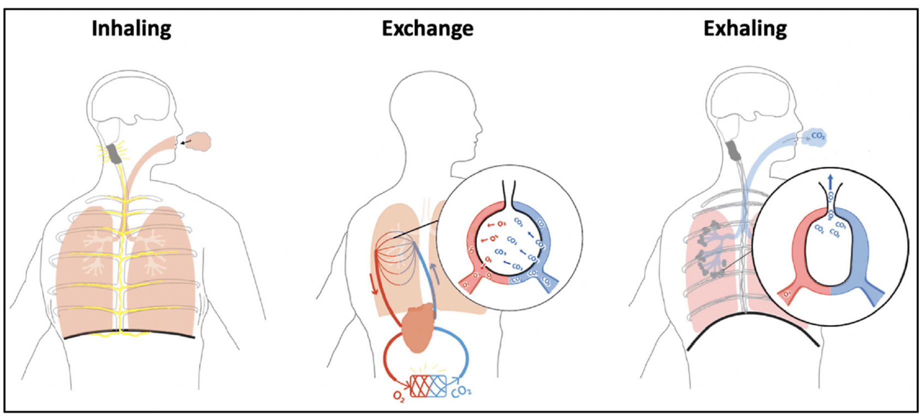

FIGURE 1 Pictures provided to the provided drawing-then-explain group during the learning activity

The spatial ability test was the paper folding test (Ekstrom et al., 1976), which consisted of one sheet of paper with instructions on one side and 10 problems on the other side. For each problem, students were asked to imagine what a piece of would look like after being folded, punched with a hole, and then reopened. Prior research suggests the Paper Folding Test is reliable (e.g., Ekstrom et al., 1976; Zhang & Fiorella, 2021, α's > .70) and performance is positively associated with learning in biology (Castro-Alonso & Uttal, 2019; Fiorella & Mayer, 2017).

2.3 | Procedure

Participants were assigned randomly to one of the four conditions: draw-and-explain, explain-and-draw, provided explanation-then-draw, or provided drawing-then-explain. Each participant completed the session individually at a laboratory computer station. All participants were aware at the beginning that the study consisted of two parts separated by 1 week. After providing informed consent, participants completed the participant survey. Then the researcher provided instructions for the learning phase of the experiment. They were told that in Part 1 they would study a lesson on the human respiratory system for 20 min and complete a brief learning activity, and that in Part 2 (1 week later) they would be tested on their understanding of the respiratory system.

2.3.1 | Part 1

After studying the lesson, participants completed the first learning experience survey, which assessed their self-reported mental effort while studying the lesson, perceived difficulty of the lesson, helpfulness of the lesson, and their perceived level of understanding of the respiratory system. Next, participants received instructions based on their assigned learning activity. All students received a total of 10 min to complete the learning activity. Students in the draw-then-explain and explain-then-draw conditions received 5 min for writing an explanation and 5 min for drawing pictures. Students in the provided drawing-then-explain and provided explanation-then-draw conditions had 5 min to study the provided text or pictures and then 5 min to either draw pictures based on the provided text or write explanations based on the provided pictures.

Following the learning activity, the experimenter directed participants to complete the second learning experience survey, which assessed their mental effort during the learning activity, perceived difficulty of the learning activity, helpfulness of the learning activity, and their perceived level of understanding of the respiratory system. Participants were then thanked for their participation in Part 1 of the study and reminded to return to the laboratory in 1 week to complete Part 2. The total duration for Part 1 was approximately 45 min.

2.3.2 | Part 2

Participants returned to the laboratory 1 week later to complete the post-test and the third learning experience survey. Students were given 5 min to complete the explanation test, 4 min for each of the two questions on the drawing test, and 3 min for each of the five questions on the transfer test. Participants were encouraged to use the total allotted time for each question and were asked not to move to new questions until the time had elapsed. After the post-test, students completed the third learning experience survey, which assessed their mental effort while completing the post-test, their perceived difficulty of the post-test, their perceived level understanding of the respiratory system, and how helpful they found the learning activity from the previous week. Finally, students were thanked for their participation in the study and awarded course credit. The total duration of Part 2 was approximately 30 min.

3 | RESULTS

3.1 | Preliminary analyses

We first tested whether the groups differed on background characteristics. One-way ANOVA indicated no significant differences across groups in self-reported prior knowledge, $F(3, 144) = 0.19$, $p = .904$, confidence in learning about science, $F(3, 144) = 0.76$, $p = .518$, interest in learning about science, $F(3, 144) = 0.44$, $p = .722$, or performance on the spatial ability test, $F(3, 143) = 2.09$, $p = .105$.

3.2 | Learning activity and post-test performance

Table 1 presents correlations among scores on the learning activities (explanation and drawing quality) and scores on the delayed post-tests (explanation test, drawing test, and transfer test). As shown in the table, each of the learning activity measures were positively correlated with each of the post-test measures, and the post-test measures were all positively correlated with each other.

Table 2 presents the means and SD for performance on the learning activity and the post-tests administered 1 week following the learning activity. To test our primary hypotheses, we used one-way ANOVA[2] with four planned contrasts: (a) Sequence hypothesis 1 (draw-then-explain > explain-then-draw), (b) Sequence hypothesis 2 (provided drawing-then-explain > provided explanation-then-draw), (c) Scaffolding hypothesis 1 (provided drawing-then-explain > draw-then-explain), and (d) Scaffolding hypothesis 2 (provided explanation-then-draw > explain-then-draw).

3.2.1 | Sequence hypothesis 1: Draw-then-explain > explain-then-draw

We did not find support for sequence hypothesis 1. The draw-then-explain and explain-then-draw groups not differ significantly in explanation quality during the learning activity, $t(109) = 1.55$, $p = .124$, or on the delayed post-tests: explanation test, $t(141) = 1.63$, $p = .106$,

TABLE 1 Correlations among learning activities and post-test performance

	Explanation LA	Drawing LA	Explanation test	Drawing test
Explanation LA				
Drawing LA	.59*			
Explanation Test	.70*	.38*		
Drawing Test	.56*	.44*	.64*	
Transfer Test	.38*	.26*	.44*	.43*

Abbreviations: LA, learning activity; Test, delayed post-test.
*$p < .01$.

TABLE 2 Descriptive statistics across groups for learning activity and post-tests

	Learning activity				Delayed post-tests					
	Explanation		Drawing		Explanation		Drawing		Transfer	
Group	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Draw-then-explain	9.0	4.6	3.9	3.0	6.4	3.8	6.7	3.8	4.4	1.8
Explain-then-draw	7.4	3.7	3.0	2.7	5.1	3.4	6.3	4.8	3.9	1.9
Provided drawing-then-explain	12.6	5.2	—	—	7.5	4.3	7.2	4.1	4.7	1.8
Provided explanation-then-draw	—	—	7.0	3.8	4.9	2.8	7.1	3.9	4.6	2.3

drawing test, $t(141) = 0.45$, $p = .657$, transfer test, $t(141) = 1.02$, $p = .308$.

3.2.2 | Sequence hypothesis 2: Provided drawing-then-explain > provided explanation-then-draw

We found partial support sequence hypothesis 2. The provided drawing-then-explain group performed significantly higher than the provided explanation-then-draw-group on the explanation post-test, $t(141) = 2.84$, $p = .005$, $d = .71$, but not on the drawing post-test, $t(141) = 0.12$, $p = .907$, or the transfer post-test, $t(141) = 0.34$, $p = .733$.

3.2.3 | Scaffolding hypothesis 1: Provided drawing-then-explain > draw-then-explain

We found partial support for scaffolding hypothesis 1. The provided drawing-then-explain group generated significantly higher-quality explanations during the learning activity than the draw-then-explain group, $t(109) = 3.44$, $p = .001$, $d = .73$. However, the two groups did not significantly differ on the post-tests administered 1 week later: explanation test, $t(141) = 1.29$, $p = .200$, drawing test, $t(141) = 0.47$, $p = .639$, transfer test, $t(141) = 0.76$, $p = .450$.

Next, we conducted a mediated regression with bootstrapping, following guidelines from Hayes and Preacher (2014), to investigate whether explanation quality mediated the relationship between learning activity and post-test performance. The total effect of learning activity on post-test performance was not statistically significant for any of the post-tests: explanation test ($b = .35$, SE $= .30$, $t[77] = 1.17$, $p = .246$), drawing test ($b = .15$, SE $= .30$, $t[77] = 0.50$, $p = .620$), or transfer test ($b = .11$, $SE = .14$, $t[77] = 0.81$, $p = .418$). However, the indirect effect was statistically significant for all three post-tests: explanation test ($b = .68$, $SE = .23$, 95% CI [.24, 1.14]), drawing test ($b = .54$, $SE = .19$, 95% CI [.19, .93]), and transfer test ($b = .13$, $SE = .06$, 95% CI [.03, .26]). Thus, the mediation effect of explanation quality between learning activity and post-test performance is indirect only. Mediation analysis diagrams for scaffolding hypothesis 1 are shown in Appendix A.

3.2.4 | Scaffolding hypothesis 2: Provided explanation-then-draw > explain-then-draw

We found partial support for scaffolding hypothesis 2. The provided explanation-then-draw group generated significantly higher-quality drawings during the learning activity than the explain-then-draw group, $t(108) = 5.14$, $p < .001$, $d = 1.22$. However, the two groups did not significantly differ on the delayed post-tests: explanation test, $t(141) = 0.17$, $p = .864$, drawing test, $t(141) = 0.73$, $p = .469$, transfer test, $t(141) = 1.30$, $p = .237$.

Next, we conducted a mediated regression analysis to investigate whether drawing quality mediated the relationship between learning activity and post-test performance. The total effect of learning activity on post-test performance was not statistically significant for any of the post-tests: explanation test ($b = -.12$, $SE = .78$, $t[63] = -0.16$, $p = .874$), drawing test ($b = .73$, $SE = 1.10$, $t[63] = 0.66$, $p = .509$), and transfer test ($b = .59$, $SE = .52$, $t[63] = 1.13$, $p = .262$). The indirect effect was statistically significant for the explanation test (b = 1.52, SE = .56, 95% CI [.60, 2.80]), and the drawing test (b = 2.37, SE = .83, 95% CI [1.03, 4.30], but not the transfer test (b = .58, SE = .32, 95% CI [−.05, 1.22]. Thus, the mediation effect of drawing quality between learning activity and explanation and drawing test performance in indirect only. Mediation analysis diagrams for scaffolding hypothesis 2 are shown in Appendix B.

3.3 | Learning experience survey

Finally, we explored whether groups differed in their responses on the learning experience surveys. The only significant differences occurred for students' perceived difficulty of the post-test, $F(3, 141) = 3.73$, $p = .013$. Students in the provided explanation-then-draw group perceived the post-test to be less difficult than students in the provided drawing-then-explain group, $t(141) = 3.01$, $p = .003$, and the explain-then-draw group, $t(141) = 2.60$, $p = .010$. There were no other significant differences across groups for ratings of effort, difficulty, helpfulness, or level of understanding (F's < 1.2, p's > .35).

4 | DISCUSSION

Prior research suggests explaining and drawing are effective strategies for learning in science. There is also some evidence that these strategies complement each other, analogous to how the integration of instructor-provided explanations and visuals supports understanding. The present study tested the idea that visuals (either provided or generated) uniquely serve to help students explain what they are learning to themselves, resulting in better understanding. Based on this view, we expected translating visuals into explanations would support understanding better than translating explanations into drawings (sequence hypothesis). Furthermore, we also expected students would benefit more from using provided explanations or visuals (rather than their own) to scaffold this translation process (scaffolding hypothesis).

Results provided limited support for the sequence hypothesis. First, students who created drawings before explaining did not perform higher (on the learning activity or post-tests) than students who wrote explanations before drawing (sequence hypothesis 1). This indicates that when explanations and drawings are both learner-generated, one sequence was not more effective than the other. Second, students who used provided drawings to write an explanation performed higher on the explanation post-test than students who used provided explanations to create drawings, though they two groups did not differ in performance on the drawing test or the transfer test (sequence hypothesis 2). One potential limiting factor is that the transfer test had relatively low reliability, though performance on the transfer test was positively correlated with all other learning activity and post-test measures (see Table 1). The selective effect on the explanation post-test might simply be due to students in the explain-from-pictures group receiving practice in generating an explanation during the learning activity, whereas the draw-from-explanations group did not generate an explanation during the learning activity.

The limited evidence for sequence effects might be because studying visuals or creating drawings cannot be easily separated from explaining. Prior research in which students think aloud or explain to others while creating or studying drawings suggests students may be engaged in self-explanation processes internally *while* studying provided drawings or *while* constructing their own drawings (see Butcher, 2006; Cox, 1999; Fiorella & Kuhlmann, 2020). If so, writing explanations after studying or constructing visuals might have been redundant, and the intended differences among the experimental conditions may have been blurred. On a related note, students might have benefited more from generating oral (rather than written) explanations, as oral explanations can be generated simultaneously with studying or creating drawings (e.g., Fiorella & Kuhlmann, 2020), and they tend to be more elaborative (Lachner et al., 2018). Thus, the complementary relationship between visuals and explaining might be most pronounced when the activities are explicitly integrated. This is analogous to research on the temporal contiguity principle in multimedia learning (see Mayer & Fiorella, 2014), which suggests simultaneous presentation of words and visuals is generally more effective than sequential presentation. Future research should examine the dynamic interplay between processing or constructing visuals and real-time oral explanations to uncover precisely how drawings facilitate explaining (or vice versa).

Results partially supported our two scaffolding hypotheses. First, students generated better explanations during the learning activity when they used instructor-provided (rather than their own) drawings (scaffolding hypothesis 1). Similarly, students generated better quality drawings during the learning activity when they used instructor-provided (rather than their own) explanations (scaffolding hypothesis 2). This is consistent with prior research showing that students need support to construct accurate explanations and drawings, and with research on learning strategies more broadly (e.g., Pressley et al., 1982). Interestingly, however, receiving this form of scaffolding during the learning activity did not result in better performance on the post-test administered 1 week later. Meditation analyses suggested explanation quality (scaffolding hypothesis 1) and drawing quality (scaffolding hypothesis 2) only indirectly mediated the relationship between learning activity and post-test performance. One possibility is that the scaffolds actually served as a crutch – they helped students generate better explanations and drawings when the scaffold was available (yielding effect sizes of .73 and 1.22, respectively), but once removed for the delayed post-tests, there was no advantage compared to the groups who generated their own explanations *and* drawings during the learning activity. This finding highlights the importance of using delayed post-test measures in educational research, as interventions that work in the short-term do not necessarily persist over time. Future research should explore additional forms of instructional support that might better support persistent knowledge. For example, rather than scaffolding, students might benefit from generating their own explanations and drawings first, and then receiving feedback and an opportunity to revise their generated representations (e.g., Zhang & Fiorella, 2019, 2021).

An alternative interpretation of this study's findings is that the learning activity primarily served as retrieval practice (Karpicke, 2012) rather than as *generative* activities that encourage students to actively organize and integrate the material with their existing knowledge. For the learning activity, students were asked to explain and/or draw specific parts of the lesson, and the provided explanations or drawings as scaffolds were taken directly from the lesson. It is important to note that explanation and drawing quality during the learning activity were positively correlated with transfer test performance, suggesting some alignment between the activities and generative learning outcomes. Nonetheless, the null effects across conditions on the transfer test suggest the activities could likely be improved to better support deep understanding. Future research should test explanation and drawing tasks that require learners to go beyond the specific content of the lesson, such as explaining how a concept applies in a new situation, or drawing elements of the learning material that were not explicitly depicted in the learning materials.

Finally, a potential criticism of this study is that it did not include control conditions who only explained, only created drawings, or who only studied the lesson without engaging in a generative activity. Our goal was not to test whether individual strategies were more effective than weak control groups, which is already well established in the literature. Our goal was to test specific hypotheses regarding the relationships between explaining and drawing activities–in particular the role of sequencing and scaffolding. As others have argued (Chi & Wylie, 2014; Fiorella & Zhang, 2018), these types of stronger experimental comparisons are important for advancing our understanding of how generative activities work and when they work best. Findings from the present study suggest that the sequence of explaining and drawing activities may not greatly affect learning outcomes and that the benefits of using provided representations as scaffolds may be limited. This work provides a path for future research to explore alternative ways to integrate explaining and drawing activities in science learning.

ACKNOWLEDGMENT

This research as supported by a grant from the National Science Foundation (1955348) and a National Academy of Education/Spencer Postdoctoral Fellowship awarded to Logan Fiorella.

ENDNOTES

[1] The sequence hypothesis in the present study is distinct from prior research testing the sequential presentation of provided text and pictures (e.g., Eitel & Scheiter, 2015). That research generally involves the *initial presentation* of text and pictures during learning. In the present study, all students initially study the same lesson, which involves the simultaneous presentation of text and pictures. Then participants engaged in one of the different implementations of drawing/explaining activities described in "The Current Study" section.

[2] The overall ANOVAs were statistically significant for explanation quality during the learning activity $F(2,109) = 11.33$, $p < .001$, drawing quality during the learning activity $F(2, 109) = 14.86$, $p < .001$, and explanation post-test score, $F(3, 141) = 3.70$, $p = .013$, but not for drawing post-test score, $F(3, 141) = 0.29$, $p = .832$, or transfer post-test score, $F(3, 141) = 1.02$, $p = .387$.

DATA AVAILABILITY STATEMENT

Data will be made available upon reasonable request to the corresponding author.

ORCID

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

REFERENCES

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

Ainsworth, S., & Scheiter, K. (2021). Learning by drawing visual representations: Potential, purposes, and practical implications. *Current Directions in Psychological Science.*, *30*, 61–67.

Berthold, K., Eysink, T. H., & Renkl, A. (2009). Assisting self-explanation prompts are more effective than open prompts when learning with 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.

Bobek, E., & Tversky, B. (2016). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology*, *98*, 182–197.

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

Castro-Alonso, J., & Uttal, D. H. (2019). Spatial ability for university biology education. In S. Nazir (Ed.), *Advances in human factors in training, education, and learning sciences* (pp. 283–291). Springer.

Chi, M. T. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 161–238). Lawrence Erlbaum Associates, Inc..

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., 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. (2015). Picture or text first? Explaining sequence effects when learning with pictures and text. *Educational Psychology Review*, *27*(1), 153–180.

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.

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.

Fiorella, L., & Kuhlmann, S. (2020). Creating drawings enhances learning by teaching. *Journal of Educational Psychology*, *112*(4), 811–822.

Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review*, *28*, 717–741.

Fiorella, L., & Mayer, R. E. (2017). Spontaneous spatial strategy use in learning from scientific text. *Contemporary Educational Psychology*, *49*, 66–79.

Fiorella, L., & Zhang, Q. (2018). Drawing boundary conditions for learning by drawing. *Educational Psychology Review*, *30*, 1115–1137.

Hayes, A. F., & Preacher, K. J. (2014). Statistical mediation analysis with a multicategorical independent variable. *The British Journal of Mathematical and Statistical Psychology*, *67*(3), 451–470.

Karpicke, J. D. (2012). Retrieval-based learning: Active retrieval promotes meaningful learning. *Current Directions in Psychological Science*, *21*(3), 157–163.

Lachner, A., Ly, K.-T., & Knuckles, M. (2018). Providing written or oral explanations? Different effects of the modality of explaining on students' conceptual learning and transfer. *Journal of Experimental Education*, *86*, 344–361.

Leopold, C., & Leutner, D. (2012). Science text comprehension: Drawing, main idea selection, and summarizing as learning strategies. *Learning and Instruction*, *22*, 16–26.

Mayer, R. E. (2020). *Multimedia learning* (3rd ed.). Cambridge University Press.

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

Pressley, M., Levin, J. R., & Delaney, H. D. (1982). The mnemonic keyword method. *Review of Educational Research*, *52*(1), 61–91.

Scheiter, K., Schleinschok, K., & Ainsworth, S. (2017). Why sketching may aid learning from science texts: Contrasting sketching with written explanations. *Topics in Cognitive Science*, *9*, 866–882.

Stenning, K., & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science*, *19*, 97–140.

Van Meter, P. (2001). Drawing construction as a strategy for learning from text. *Journal of Educational Psychology*, *93*, 129–140.

Van Meter, P., & Firetto, C. M. (2013). Cognitive model of drawing construction. In G. Schraw, M. T. McCrudden, & D. Robinson (Eds.), *Learning through visual displays* (pp. 247–280). Information Age Publishing, Inc..

Van Meter, P., & Garner, J. (2005). The promise and practice of learner-generated drawing: Literature review and synthesis. *Educational Psychology Review*, *17*, 285–325.

Wylie, R., & Chi, M. T. H. (2014). The self-explanation principle in multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (Second ed., pp. 413–432). Cambridge University Press.

Zhang, Q., & Fiorella, L. (2019). Role of generated and provided visuals in supporting learning from scientific text. *Contemporary Educational Psychology*, *59*, 101808.

Zhang, Q., & Fiorella, L. (2021). Learning by drawing: When is it worth the time and effort? *Contemporary Educational Psychology*, *66*, 101990.

How to cite this article: Fiorella, L., Pyres, M., & Hebert, R. (2021). Explaining and drawing activities for learning from multimedia: The role of sequencing and scaffolding. *Applied Cognitive Psychology*, 1–11. https://doi.org/10.1002/acp.3871

APPENDIX A1

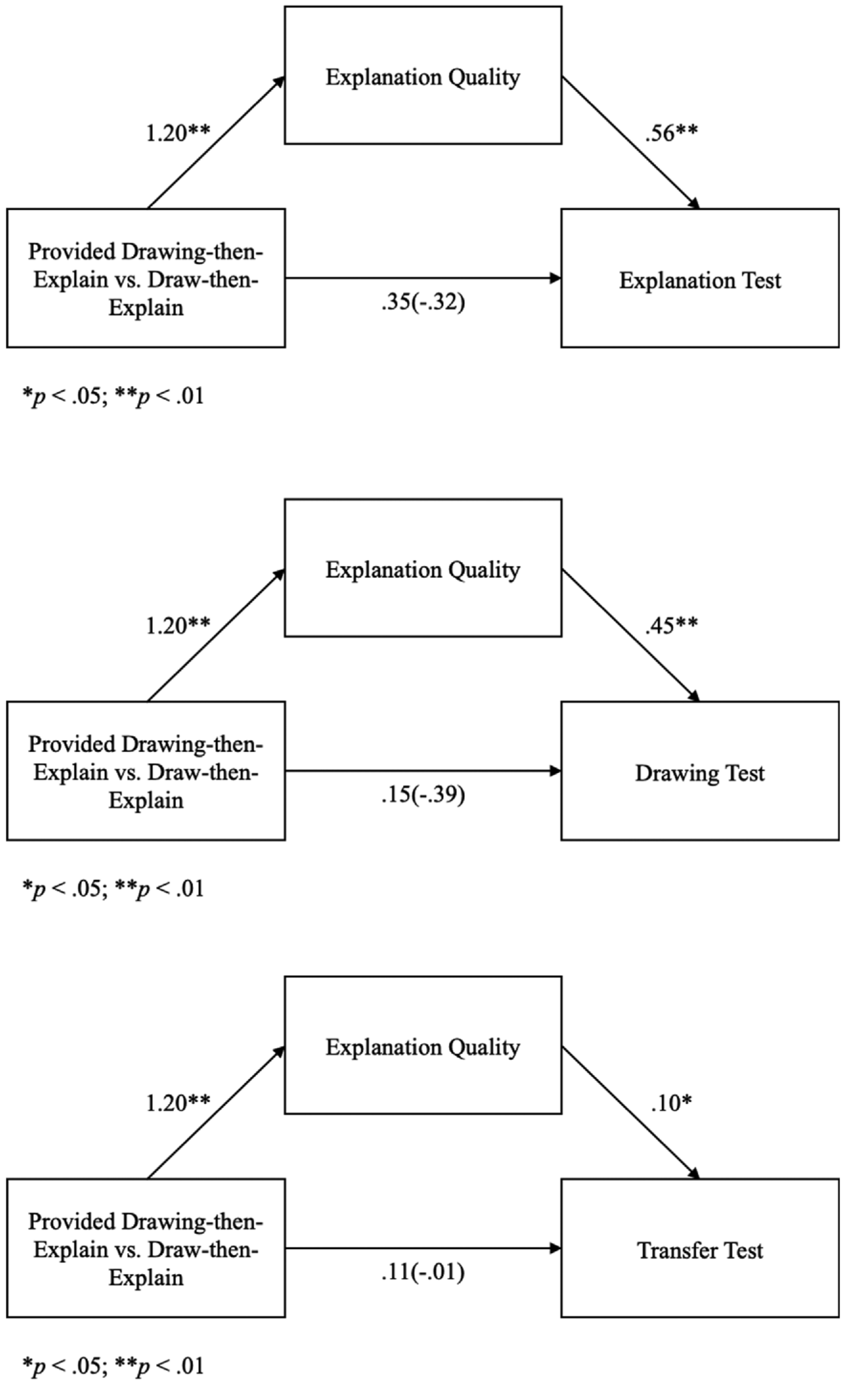

FIGURE A1 Mediation analysis diagrams for scaffolding hypothesis 1

APPENDIX B1

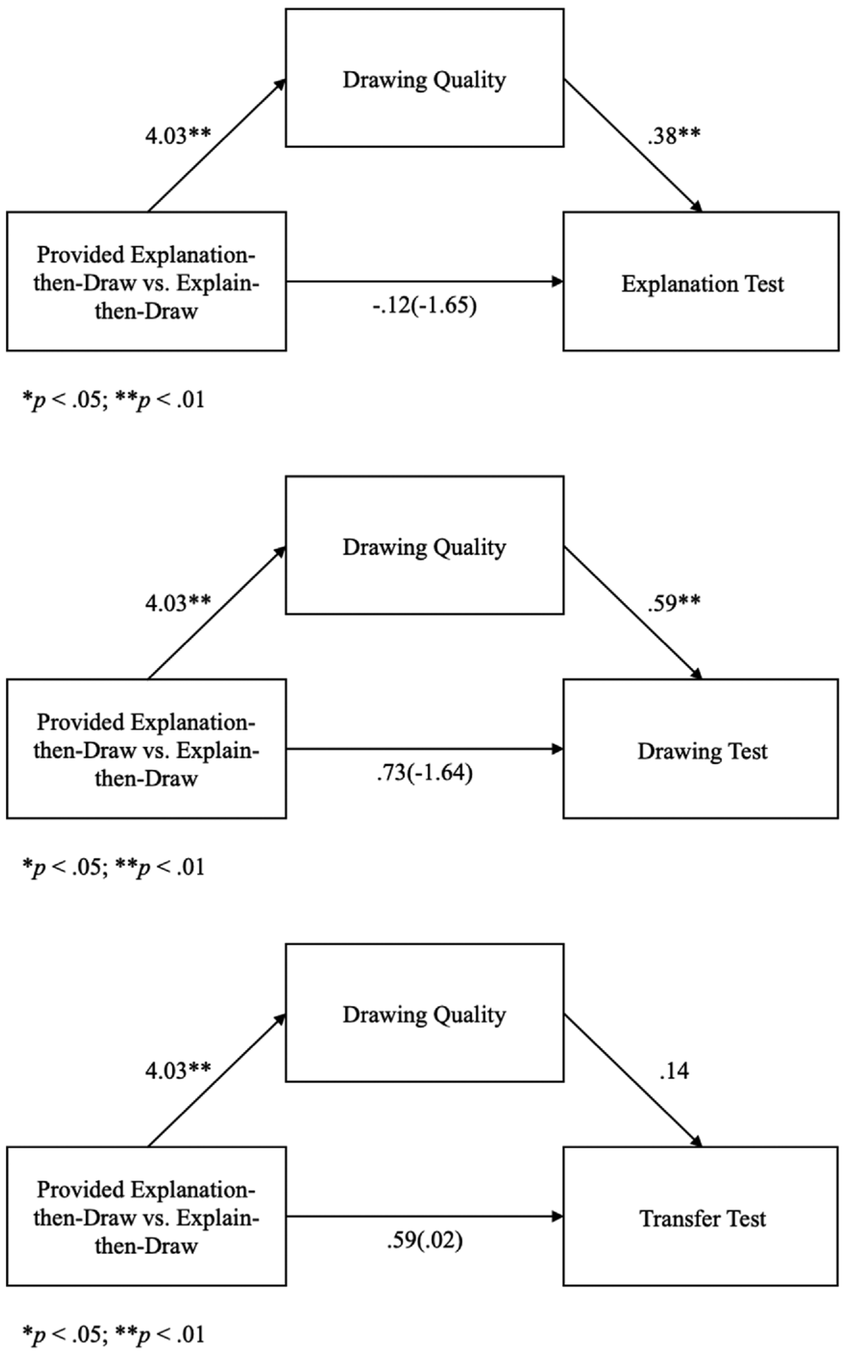

FIGURE B1 Mediation analysis diagrams for scaffolding hypothesis 2