Learning by Drawing: When is it Worth the Time and Effort?

Date Submitted: July 13, 2020

Revision Submitted: January 8, 2021

Revision 2 Submitted: May 7, 2021

Abstract

Two experiments compared the effects of learning by drawing to studying instructor-provided visuals on learning outcomes, learning time, and cognitive load. College students studied a text on the human circulatory system and completed comprehension and transfer tests. In Experiment 1 ($N = 107$), students studied the text with provided visuals (provided visuals) or generated their own drawings from a text with text-based support (verbally-supported drawing) or without support (unsupported drawing). Results showed that while the verbally-supported drawing condition spent significantly more time and experienced significantly higher cognitive load than the provided visuals condition, there were no differences across the three conditions in learning outcomes. In Experiment 2 ($N = 85$), students studied the text with provided visuals (provided visuals) or generated drawings from the text with provided visuals as feedback (visually-supported drawing). Results showed that the visually-supported drawing condition spent significantly more time and experienced significantly higher cognitive load than the studying provided visuals condition but also performed significantly better than the provided visuals condition on the comprehension test. These findings suggest generating drawings prior to studying provided visuals is worth the time and effort.

Keywords: drawing; generative activities; multimedia; cognitive load; learning time

Learning by Drawing: When is it Worth the Time and Effort?

Students generally learn better from scientific text when they are provided with relevant visuals (i.e. the multimedia learning effect; Mayer, 2014) or when they are asked to generate visuals depicting the text (i.e. the generative drawing effect; Van Meter & Firetto, 2013). In prior research, support for the multimedia learning effect and the generative drawing effect is based on comparisons to students who only learn from text or use text-based strategies *without* visualizations (Authors, 2018; Mayer, 2014). However, we know little about the relative effectiveness of different ways to learn *with* visualizations. In particular, the few studies comparing effects of creating drawings to studying provided visuals have shown mixed effects on learning outcomes. Moreover, research has rarely compared the effects of these strategies on learning time and cognitive load. The present study addresses this gap by comparing the effects of learner-generated drawings to studying instructor-provided visuals on learning outcomes, learning time, and cognitive load.

Creating Drawings vs. Studying Provided Visuals

Learning by drawing and studying provided visuals both have unique strengths and weaknesses (Authors, 2019). According to the cognitive model of drawing construction, there are two main theoretical advantages of drawing over studying provided visuals (CMDC, Van Meter & Firetto, 2013). First, drawing improves learning from text by forcing students to integrate verbal and nonverbal representations (Van Meter & Firetto, 2013; Van Meter & Garner, 2005). When students are asked to generate drawings from texts, they must use their prior knowledge to convert the text into a pictorial display. This translation process encourages students to actively engage in multimodal processing, supporting deeper understanding and knowledge transfer (Authors, 2016, 2019; Leutner & Schmeck, 2014; Stern, Aprea, & Ebner, 2003). In contrast, when students learn from provided visuals, they often process the visuals passively or they have difficulties integrating the text and corresponding visuals (De Jong et al., 1998). Eye-tracking research also shows that processing of multimedia lessons is heavily text-focused, indicating that learners do not spontaneously integrate the two representations (Renkl & Scheiter, 2017; Schnotz & Wagner, 2018). The second potential advantage of drawing is that it may enhance students' monitoring and regulation behaviors by providing students with metacognitive cues, such as experiencing difficulty while generating drawings (Van Meter, 2001). In contrast, the presence of provided visuals may induce an illusion of understanding and harm students' metacognitive judgments of learning (Peeck, 1993; Wiley, 2018).

There are also potential advantages of studying instructor-provided visuals over learning by drawing. According to the cognitive theory of multimedia learning (CTML; Mayer, 2014) and the integrated model of text and picture comprehension (ITPC; Schnotz, 2014), well-designed visuals could help students construct an *accurate* mental model from the text since the structure of external representations (i.e. the provided visuals) influences the structure of students' mental models. For instance, labeled and annotated illustrations can make referential connections between texts and illustrations explicit to facilitate the integration of verbal and nonverbal representations (Leopold et al., 2015; Mayer, 1989; Mayer et al., 1995). In contrast, the accuracy of students' mental models constructed from learning by drawing depends on drawing quality. The quality of students' drawings without sufficient instructional support is generally low, which may result in students building *inaccurate* mental models (Authors, 2018, 2019; Peeck, 1993). Another advantage of studying provided visuals is that they may allow students to achieve similar or even superior learning outcomes while requiring less time and cognitive load (Skuballa, Dammert, & Renkl, 2018; Rasch & Schnotz, 2009). Cognitive load is defined as the amount of working memory resources required for performing a learning task. Specifically, intrinsic and germane load results from cognitive activities that contribute to learning while extraneous load results from cognitive activities that are irrelevant and unnecessary for learning (Sweller, 2010). Studying provided visuals may scaffold accurate mental model construction and free up working memory resources for inference making, which is especially beneficial for students with low prior knowledge (Glenberg & Langston, 1992; Carney & Levin, 2002; Ollerenshaw et al., 1997). In contrast, drawing can be cognitively demanding and time-consuming (Hellenbrand et al., 2019; Leutner et al., 2009; Peeck, 1993; Schmeck et al., 2014; Schmidgall et al., 2019), especially for novice learners who are more easily overloaded by performing complex behavioral and cognitive activities like drawing (Seufert, 2003, 2019; Skuballa et al., 2018).

In summary, learning by drawing and studying provided visuals both have theoretical strengths and weaknesses. As noted in a recent review by Authors (2018), it is important for researchers to clarify which strategy is most effective under which conditions. The present study focuses on how generating drawings and studying provided visuals differ in their effects under different conditions not only on learning outcomes but also on learning time and cognitive load, which are important contextual factors to consider when evaluating a learning strategy (Dunlosky et al., 2013).

Empirical Research on Drawing vs. Studying Provided Visuals

The few empirical studies that have compared drawing to studying provided visuals on learning outcomes have yielded mixed findings. For example, Van Meter and colleagues (Van Meter, 2001; Van Meter et al., 2006) found superior effects of drawing over provided visuals, particularly when students revised their drawings based on provided visuals. In contrast, findings from Schwamborn et al. (2011) and Leopold et al. (2013) favored studying provided visuals because drawing activities created extraneous cognitive load. Finally, recent studies by Schmidgall et al. (2019) found no differences between the two strategies, and Hellenbrand et al. (2019) found drawing was more effective than studying provided visuals for retention but not transfer.

The mixed results might be due to different levels of drawing support implemented in these studies, given that instructional support is often necessary for students to benefit from drawing (Van Meter & Firetto, 2013). Authors (2018) identified 4 distinct levels of support from literature: minimal guidance, drawing training, providing partially completed visuals, and comparing one's drawing to provided visuals. The level of drawing support may moderate the effectiveness of drawing when compared to provided visuals. With lower levels of drawing support (e.g. Leopold et al., 2013; Schwamborn et al., 2011), drawing may be *less effective* than studying provided visuals; with higher levels of drawing support (e.g. Van Meter, 2001; Van Meter et al., 2006), drawing may be *more effective* than studying provided visuals. However, it should be noted that the benefits of higher levels of drawing support are confounded because the support used in prior research often *includes* provided visuals (e.g., providing partially completed visuals or comparing one's drawing to provided visuals). In such cases, students likely benefit from drawing *and* studying provided visuals (Authors, 2019).

Although many studies report that drawing is time-consuming and cognitively demanding, few studies have directly compared the effects of drawing to studying provided visuals on learning time and cognitive load. Past research suggests drawing may create greater cognitive load compared to studying provided visuals, impairing learning outcomes (Leutner, Leopold, & Sumfleth, 2009; Schwamborn et al., 2011). Furthermore, the potential benefits of drawing may disappear when accounting for differences in learning time (Hellenbrand et al., 2019; Schmeck et al., 2014; Van Meter, 2001). Taken together, drawing may not be worth additional time and cognitive load under certain conditions compared to studying provided visuals. For example, drawing may not be effective for novice learners who do not receive strong forms of drawing guidance. Examining time and cognitive load during learning is important because an effective strategy should enable students to invest their time and effort wisely. If a particular drawing activity takes more time and effort but leads to similar learning outcomes, this provides insight into how different types of learning with visuals contribute to learning and how students and instructors should apply generative learning activities.

In the present study, novice college students learned from a text on how the human circulatory works by creating a drawing or studying provided visuals. We assessed learning outcomes via comprehension and transfer tests, and we measured learning time and students' self-reported cognitive load. We also examined the potential moderating role of different forms of drawing support, given that the effectiveness of drawing may depend on students' ability to generate quality drawings with appropriate support.

Experiment 1

In Experiment 1, students either studied a text with provided visuals (provided visuals condition), created a drawing from the text without support (unsupported drawing condition), or created a drawing from the text with text-based support (verbally-supported drawing condition). Unlike past research, we provided a high level of drawing support that does not involve any provided visuals (e.g., partially provided visuals or comparing one's drawing to a provided visual). This is important because it allows us to isolate the effects of drawing without introducing the potential benefits of allowing students to view provided visuals (see Authors, 2018). In particular, the verbally-supported drawing condition first received textual labels to help them focus on the critical elements from the text (instead of using partially provided visuals, as in Leutner & Schmeck, 2014 and Schwamborn et al., 2010; Kollmer et al., 2020). After drawing, students received text-based questions to scaffold their revision process (instead of comparing their drawing to a provided visual, as in Van Meter, 2001 and Van Meter et al., 2006). Taken together, these two modified forms of drawing support were designed to help students generate higher-quality drawings without the effects of viewing any elements of instructor-provided visuals.

Predictions

First, we predicted the verbally-supported drawing condition would outperform the unsupported drawing condition and the provided visuals condition on learning outcomes (*learning outcomes hypothesis*). We did not have an a priori prediction regarding learning outcomes between the unsupported drawing condition and the provided visuals condition because each has theoretical strengths and weaknesses, and prior research has yielded mixed results. Second, we predicted the drawing conditions would spend significantly more time and experience a significantly higher cognitive load than the provided visuals condition (*learning time and cognitive load hypothesis*). This prediction is based on prior research showing that drawing is cognitively demanding and time-consuming. Finally, we predicted the verbally-supported drawing condition would create higher-quality drawings than the unsupported drawing condition and drawing quality would be positively associated with learning outcomes (*drawing quality hypothesis*).

Method

Participants and design

Participants were 107 undergraduates recruited from a large university in the southeast United States who received course credit (*n* = 44) or gift cards (*n* = 63) for participation. 61% were non-STEM majors and 66% were freshmen or sophomores. The mean age was 19.89 (*SD* = 1.33), and there were 33 men and 74 women. Performance on a prior knowledge test (see below) indicates this population has relatively low prior knowledge related to the circulatory system. Participants were randomly assigned to one of the three conditions: provided visuals (*n* = 34), unsupported drawing (*n* = 37), and verbally supported drawing (*n* = 36). This study was approved by and conducted in accordance with the ethical standards of the authors' Institutional Review Board (IRB).

Materials

Materials consisted of a prior knowledge test, a text on the circulatory system, provided visuals, drawing papers, comprehension and transfer tests, a cognitive load questionnaire, a demographic survey (e.g., age, gender, major), and a measure of spatial ability test. The prior knowledge test consisted of 10 short-answer questions (α = .75) asking specific facts about the anatomy and function of the human circulatory system (adapted from Wolfe et al., 1998). Example questions include "Blood returning from the body enters which chamber of the heart first?" and "How many continuous, closed circuits of blood are there from the heart? Name them." Answers were coded by awarding one point for each correct response; some questions were worth more than one point. For example, answers to the second question include: two circuits; and the pulmonary and systemic circuits. For this question, one point was awarded for the correct number of circuits and one point for each name of the two circuits, resulting in three possible points. Such coding yielded a total of 24 possible points. The prior knowledge test was scored by one research assistant since each question required only a 1- or 2-word objective response.

The text lesson was a two-page, paper-based lesson containing 1108 words and 9 paragraphs. The lesson explained how the human circulatory system works by describing the four heart chambers, the major vessels, the four heart valves, and the structural and functional relations among them (adapted from Marieb & Hoehn, 2016). According to the Flesh-Kincaid readability index, the text is written at the 10th-grade level. The provided visuals were a one-page, black and white diagram representing the content of the text (from Reece et al., 2010), including labels and icons of the heart chambers, vessels and the heart valves, and arrows to show the paths of blood flow. The text lesson and diagram of the circulatory system are presented in Appendix A.

The drawing paper for the unsupported drawing condition was one sheet of 8.5 × 11 paper with general instructions at the top asking students to draw a picture or multiple pictures on the drawing paper to help them understand how the circulatory system works (see Appendix B). For the verbally-supported drawing condition, the same size drawing paper included instructions at the top, a simple human body background in the middle, and textual labels of the most important elements in the circulatory system at the bottom to help students focus on critical elements in the system (see Appendix B). The instructions asked students to draw a picture by using the labels at the bottom to name different parts. Students were asked to use all the labels and told their drawing should reflect relationships among the different parts described in the text. The instructions also included basic drawing conventions, such as using lines with arrows to represent the direction of blood flow.

After creating their drawing, students in the verbally-supported drawing condition answered a series of text-based questions to help them revise their drawings, and they were asked to judge the quality of their drawings. For example, one set of questions included "(a) Which chambers are located on the top of the heart in your drawings? (b) Which chambers are located at the bottom of the heart in your drawings? (c) Are the locations correct? Revise your drawings as necessary". Possible answers to this set of questions include (a) the left and right atrium; (b) the left and right ventricles; and (c) yes, my drawing is correct, respectively. Students answered the questions in writing based on their drawings. We coded students' judgments of their drawings (the final question in each question set) by assigning one point to students who judged their drawings to be correct and 0 points to students who judged their drawings to be incorrect. We also assigned 0.5 points to some students (n=5) who answered they were not sure about their drawings. If students did not make an explicit judgment by answering the question, we checked whether they revised their drawings. We assumed that students who revised their corresponding drawings based on the questions judged their drawings to be incorrect; if they did not revise their drawings, we assumed they judged their drawings to be correct.

We developed a coding scheme from the text to measure the quality of students' drawings (adapted from Authors, 2019), which included elements, relations, and systems. Specifically, there were 21 elements of the human circulatory systems, including the heart chambers, valves, major vessels and relevant parts in the lungs and human body. Relations consisted of 10 structural relations (e.g. the right atrium connects to the pulmonary veins) and 12 functional relations among elements (e.g. the right atrium receives O2 rich blood from the pulmonary veins). Finally, the elements and relations assemble three systems (i.e. the circulatory system, the pulmonary system, and the valves system) that ensure the human circulatory system works appropriately. Students were awarded one point for each correctly drawn element and relation. For a system, students received one point if the elements and relations in their drawings correctly reflected the corresponding system. For example, to receive one point for the pulmonary circulation system, students needed to first draw out elements such as the right ventricle, the pulmonary valve, the pulmonary trunk and arteries, the lungs, the pulmonary vein, and the left atrium. Then students should also correctly depict the relations among these elements. If students' drawings correctly reflect only parts of the system without including all the elements and relations, students received half of a point. Students received zero points if any relation among elements was incorrect. The coding scheme yielded a total of 46 possible points. Since students used different colors of pens in generating and revising drawings, we were able to code students' drawings for both study sessions separately. Two research assistants scored students' drawings, and discrepancies were settled through discussion. Intraclass correlation coefficients (ICC) between total scores of two research assistants for the first 25% of the data ($n = 28$) was calculated based on a single-rating, absolute-agreement, 2-way mixed-effects model. The results showed good inter-rater reliability, ICC = .82, 95% CI [.69, .90], so one assistant scored all remaining data.

Learning outcomes were assessed with a comprehension test and a transfer test. The comprehension test was one free-response question asking students to explain how the human circulatory works in detail based on what they learned from the lesson. The coding scheme for scoring the comprehension test was similar to that used for scoring students' drawings, which distinguished among elements, relations, and systems. Additionally, the coding scheme included individual "idea units" corresponding to specific facts described in the text. For example, students could receive points for describing that "alveoli are wrapped in a network of capillaries". Since the coding scheme was based on the text lesson only, information needed to perform well in the comprehension test was equivalent across conditions. Students received one point for each correctly explained element, relation, and system, as well as each individual idea unit included in their response, for a total of 55 possible points. ICC between total scores of two research assistants for the first 20% of the data (n = 25) was calculated based on a single-rating, absolute-agreement, 2-way mixed-effects model. The results showed good inter-rater reliability, ICC = .75, 95% CI [.60, 87], so one assistant scored all remaining data.

The transfer test consisted of 4 open-ended questions that asked students to apply what they learned from the lesson to new contexts (adapted from Authors, 2019; Butcher, 2006; Marieb & Hoehn, 2016). The average interitem correlation for the 4-item transfer test was .22, which is within the recommended range of .15 to .50 (Clark & Watson, 1995). All the individual interitem correlations also fell in the range of .15 to .50, which further suggested homogeneity of the measure (Clark & Watson, 1995). Answers to the transfer questions were not addressed explicitly in the lesson; students needed to make inferences and predictions based on their mental models of the human circulatory system. The transfer questions consisted of: (1) "What will happen to the LUNGS if ONLY the left ventricle stops pumping but the right ventricle keeps pumping? What will happen to the BODY if ONLY the right ventricle stops pumping but the left ventricle keeps pumping? Please explain your answer in detail."; (2) "Imagine a disease that causes the walls of all blood vessels in the body to grow dramatically thicker, effectively narrowing the width of all vessels. What effect do you think this would have on the HEART? Please explain your answer in detail."; (3) "Carbon monoxide (CO) binds with the hemoglobin in red blood cells better than it does with either oxygen or carbon dioxide (CO2). If someone has inhaled a great deal of CARBON MONOXIDE (CO), what will happen to your body? Please explain your answer in detail."; and (4) "What will happen to the HEART if valves cannot close tightly? Please explain your answer in detail." Students received one point for each acceptable response regarding the predictions of the situation described in the question (i.e. what will happen), and one point for each correct component of their explanation of their predictions (i.e. why it will happen). For example, acceptable responses concerning the results of the second transfer question include: (a) the left ventricle will overwork and become weaker; and (b) if untreated, it will lead to heart failure. Acceptable explanations of the results include: (a) the thickened blood vessels will decrease blood flow; (b) and cause high blood pressure; (c) so the left ventricle needs to work harder to pump against increased resistance. Thus, this transfer question was worth 5 possible points. In total, the transfer test was worth 29 possible points. ICC between total scores of two research assistants for the first 20% of the data ($n = 25$) was calculated based on a single-rating, absolute-agreement, 2-way mixed-effects model. The results showed good inter-rater reliability, ICC = .73, 95% CI [.68, .80], so one assistant scored all remaining data.

The cognitive load questionnaire contained two questions. The first question was "How difficult was the lesson you just studied?" with a 7-point subjective rating scale developed by Kalyuga, Chandler, and Sweller (1999), ranging from (1) extremely easy to (7) extremely difficult. The second question was "What level of effort did you put into learning the lesson?" with a 7-point subjective rating scale developed by Paas and Van Merriënboer (1993), ranging from (1) extremely low to (7) extremely high. We calculated the sum of the two ratings as the measure of students' overall cognitive load during learning (see Paas et al., 2003; Gog & Paas, 2008).

The spatial ability test was the Paper Folding Test (Ekstrom, Dermen, & Harman, 1976) ($\alpha = .71$), which consisted of instructions and examples on one page and 10 problems on a second page. Each problem included figures representing a square piece of paper being folded and punched with a hole, and it asked students to imagine where the hole(s) will be after reopening the paper. Past research indicates that the Paper Folding Test loads onto the Spatial Visualization sub-factor of spatial ability (Carroll, 1993), which is especially predictive of achievement in STEM disciplines (Uttal & Cohen, 2012). Other research has found that the Paper Folding Test is positively associated with science text comprehension (Authors, 2017, 2019; Castro-Alonso & Uttal, 2019). Answers were scored by tallying the number of correct responses out of 10.

Procedure

Students were randomly assigned to one of the three conditions and were tested individually in a laboratory. After students read and signed the consent form, the experimenter gave brief instructions about the study. Next, students completed the demographic survey and pretest on a computer at their own pace. Following the pretest, all students received the same general instructions for studying the circulatory system lesson: "Please carefully study the lesson explaining the human circulatory system and later you will be tested on what you learned.". Then students received specific learning strategy instructions corresponding to their assigned experimental conditions (see Appendix D for details). In the first study session, the provided visuals condition studied the text with the provided visuals of the human circulatory system without any drawing activities. The unsupported drawing condition used the drawing paper with only a general drawing instruction to create a drawing based on the text. The verbally-supported drawing condition used the drawing paper with additional labels of critical elements to create a drawing. Students worked at their own pace and let the experimenter know when they were finished. After the first study session, the experimenter collected the learning materials and drawing papers, and students completed the two cognitive load questions. Students took a two-minute break before the second study session.

In the second study session, the provided visual condition restudied the text and provided visuals. The unsupported drawing condition restudied the text and were asked to revise their drawings as necessary. The verbally-supported drawing condition answered the text-based, guided revision questions and revised their drawings as necessary by comparing their answers to the text lesson. Students were again told to work at their own pace and let the experimenter know they were finished. After the second study session, students completed the two cognitive load questions again.

The experimenter recorded the study time for each study session. After the two study sessions, students completed the self-paced comprehension and transfer tests by typing their responses on a computer. Finally, students completed the paper-based spatial ability test for 3 minutes.

Results

Preliminary Analysis

First, we tested if the three conditions differed on spatial ability and prior knowledge (see Table 1). Exploratory data analysis indicated that prior knowledge scores were highly skewed, so we used the nonparametric Kruskal-Wallis rank test to compare conditions in prior knowledge, and we used ANOVA to compare conditions on spatial ability. The conditions did not significantly differ in prior knowledge, $\chi^2(2) = .83$, $p = .661$ or spatial ability, $F(2, 104) = 1.21$, $p = .303$.

Next, we computed Pearson correlation coefficients among prior knowledge, spatial ability, learning time, cognitive load, and comprehension and transfer performance. As shown in Table 2, prior knowledge and spatial ability were significantly positively associated with transfer performance, but only prior knowledge was significantly positively associated with comprehension performance. Therefore, we considered the role of prior knowledge and spatial ability as possible moderators or covariates when analyzing transfer performance, and the role of prior knowledge as a possible moderator or covariate when analyzing comprehension performance.

Did the Conditions Perform Differently on Learning Outcomes?

For comprehension, we conducted a multiple linear regression analysis with condition, prior knowledge, and an interaction term between them as predictors in predicting comprehension. The ANOVA table showed no main effect of condition, $F(2,101) = .31$, $p = .732$, but a significant interaction effect between prior knowledge and condition, $F(2,101) = 4.13$, $p = .019$, $\eta^2 = .08$. To examine the interaction effect, we first assigned students into higher and lower prior knowledge conditions based on the median of their prior knowledge performance. Then we did a simple effect analysis. Planned contrasts showed that among higher prior knowledge students, there was no significant difference between the provided visuals condition and the drawing conditions, $t(101) =- .50$, $p = .616$. However, the unsupported drawing condition performance significantly better than verbally supported drawing condition, $t(101) = -2.21$, $p = .030$, $d = 1.02$. Among lower prior knowledge students, there was no significant difference between the provided visuals condition and the drawing conditions, $t(101) = .16$, $p = .875$, or between the unsupported drawing condition and the verbally-supported drawing condition, $t(101) = .01$, $p = .989$.

For transfer, we conducted a multiple linear regression analysis with condition, prior knowledge, spatial ability and interaction terms between them as predictors in predicting transfer. The ANOVA table showed no main effect of condition, $F(2,98) = .85$, $p = .430$ and there was no significant interaction between prior knowledge and condition, $F(2,98) = 1.33$, $p = .269$, or between spatial ability and condition, $F(2,98) = 1.53$, $p = .223$. Therefore, we conducted an ANCOVA with condition as the independent variable and prior knowledge and spatial ability as covariates. In general, there were no significant differences among conditions, $F(2, 102) = .99$, $p = .374$. Planned contrasts revealed that there was no significant difference between the provided visuals condition and the unsupported drawing condition, $t(102) = -1.3$, $p = .323$, or between the provided visuals condition and the verbally supported drawing condition, $t(102) = -.26$, $p = .94$.

These findings did not support our learning outcome hypothesis that students who learn by verbally supported drawing would perform better than students who study the provided visuals or learn by unsupported drawing. In other words, the provided visuals condition performed as well as the unsupported and verbally supported drawing condition on both comprehension and transfer.

Did the Conditions Differ in Learning Time and Cognitive Load?

We used ANOVA to compare conditions on learning time and cognitive load (averaged over two study sessions; Table 4). For learning time, there was a significant difference among conditions, $F(2, 104) = 22.65$, $p < .001$. Planned contrasts showed that the verbally-supported drawing condition spent significantly more time than the provided visuals condition, $t(104) = 6.45$, $p < .001$, $d = 1.39$. There was no significant difference between the provided visuals condition and unsupported drawing condition, $t(104) = 1.68$, $p = .173$. For cognitive load, there was a significant difference among conditions, *F*(2, 104) = 6.95, *p* = ,002. Planned contrasts showed that the verbally-supported drawing condition experienced significantly higher cognitive load than the provided visuals condition, *t*(104) = 3.46, *p* = .002, *d* = .84. There was no significant difference between the provided visuals condition and unsupported drawing condition, *t*(104) = .60, *p* = .766.

Overall, these findings partly support our learning time and cognitive load hypothesis that students who learned by drawing with verbal support spent significantly more time and experienced significantly higher cognitive load than students who studied provided visuals.

Did Instructional Support Increase Drawing Quality?

Drawing quality was significantly positively associated with comprehension ($r = .38$, $p < .001$, after study session 1; $r = .42$, $p < .001$, after study session 2) and transfer ($r = .31$, $p = .007$, after study session 1; $r = .32$, $p = .005$, after study session 2) which is consistent with past research (Schwamborn et al., 2010). Table 1 presents the means and standard deviations of drawing quality for the unsupported and verbally-supported drawing conditions. Welch's two-sample independent t-tests showed that there was no significant difference in drawing quality between the unsupported drawing condition and the supported drawing condition, *t*(48.24) = -1.54, *p* = .131. Since instructional support for the two study sessions was different, we further compared students' drawing quality after each study session. Welch's two-sample independent t-tests showed that the supported drawing condition created higher-quality drawings than the unsupported drawing condition after the first study session, *t*(51.003) = -3.34, *p* = .002, but there was no significant difference in drawing quality between the two conditions after the second study session, *t*(48.43) = -1.46, *p* = .150. This suggests that only the study session 1 support for selecting critical elements was effective but not the study session 2 support for revising drawings. These findings partially support our drawing quality hypothesis.

Why Was the Study Session 2 Drawing Support Ineffective?

As a follow-up analysis, we explored why the drawing support provided during study session 2 appeared ineffective. Students in both drawing conditions were asked to revise their drawings in study session 2, so first, we calculated the percentage of students who took the opportunity to revise their drawings. 70% ($n = 26$) of students in the unsupported drawing condition revised their drawings, but only 44% ($n = 16$) of students in the supported drawing condition did so. Moreover, paired-samples *t*-tests showed that there was a significant improvement in students' drawing quality in the unsupported drawing condition between the two study sessions, $t(36) = -7.66$, $p < .001$, but there was no significant improvement in the supported drawing condition, $t(35) = 1.71$, $p = .096$. Furthermore, students in the supported drawing condition were expected to improve their drawing quality in study session 2 by accurately judging the quality of their drawings based on the guided questions and revising their drawings accordingly. However, we found no significant correlation between students' self-judgement scores and their drawing quality after study session 1 ($r = .16$, $p = .350$). This indicates that the text-based guided revision questions did not improve students' judgement accuracy and, thus, did not help students revise their drawings.

Experiment 2

Independent of time and effort, there were no significant differences in learning outcomes across conditions in Experiment 1, suggesting that studying well-designed visuals and generating drawings are equally effective. Furthermore, students who learned by drawing with verbal support spent significantly more time and experienced significantly higher cognitive load than students who studied provided visuals. This suggests drawing activities with only verbal forms of support may not be worth the added time and effort. In addition, the verbally-supported drawing condition did not outperform the unsupported drawing condition. One explanation is that the drawing support without the benefit of viewing provided visuals (especially in study session 2) was insufficient to help students improve their drawing quality. Therefore, in Experiment 2 we implemented the highest form of drawing support (based on the classification by Authors, 2018), by asking students to compare their drawings to the provided visuals and revise drawings accordingly (e.g., Van Meter, 2001). Compared to Experiment 1, such support is even stronger since students benefit from both drawing their own visuals *and* studying provided visuals (Authors, 2019). Moreover, because prior research (and findings from Experiment 1) suggests appropriate guidance is necessary for drawing to be *more effective* than studying provided visuals, we did not include an unsupported drawing condition in Experiment 2. Thus, Experiment 2 compared the effects of studying provided visuals (provided visuals condition) to creating drawings with strong support (visually-supported drawing condition) on learning outcomes, learning time, and cognitive load. The goal was to determine whether drawing activities are worth the added time and effort when combined with an opportunity to study provided visuals. Similar to Experiment 1, we predicted the visually-supported drawing condition would outperform the provided visuals condition on learning outcomes (*learning outcomes hypothesis*) and would spend significantly more learning time and experienced significantly higher cognitive load than the provided visuals condition (*learning time and cognitive load hypothesis*).

Method

Participants and design

Eighty-five undergraduates were recruited from the Educational Psychology Subject Pool at a large university in the southeast United States who received course credit for participating. 96% were non-STEM majors and 53% were freshmen or sophomores. The mean age was 20.22 (*SD* = 1.85), and there were 17 men and 68 women. Participants were randomly assigned to one of the two conditions: provided visuals (*n* = 44) and visually supported drawing (*n* = 41). This study was approved by and conducted in accordance with the ethical standards of the authors' Institutional Review Board (IRB).

Materials

The materials were mostly identical to Experiment 1, with two exceptions. First, the drawing paper for the visually supported drawing condition contained only instructions at the top without a human body background in the middle or labels at the bottom. Second, the drawing revision questions from Experiment 1 were removed (see Appendix C for details). We made these changes because Experiment 2 aimed to investigate a different form of drawing support (i.e. compare to the provided visuals), as described below.

Procedure

The procedure was mostly identical to Experiment 1, with the exception of how participants studied the lesson. Students in the provided visuals condition studied the text with the provided visuals of the human circulatory system. Students in the visually-supported drawing condition were instructed to create a drawing of the circulatory system based on the text. After they finished drawing, the experimenter additionally provided them with the same visuals as the provided visuals condition and asked them to revise their drawings based on the visuals. They had access to the text throughout the learning phase (see Appendix E for details). Thus, unlike in Experiment 1, Experiment 2 consisted of one study session, and a different form of instructional support was provided for the visually-supported drawing condition. All other aspects of the procedure for Experiment 2 were the same as in Experiment 1.

Results

Preliminary Analysis

First, we tested if the two conditions differed on spatial ability and prior knowledge (see Table 5). As in Experiment 1, exploratory data analysis indicated that prior knowledge for both conditions was highly skewed, so we used the nonparametric Wilcoxon rank test of median to examine equivalence in conditions for prior knowledge, and we used two-sample independent *t*-tests for spatial ability. The conditions did not significantly differ in prior knowledge, $W = 1074.5$, $p = .12$, or spatial ability, $t(81.27) = .97$, $p = .335$.

Next, we computed Pearson correlation coefficients among the prior knowledge, spatial ability, learning time, cognitive load, the comprehension and transfer test. As shown in Table 6, prior knowledge was significantly positively associated with comprehension and transfer test performance. Thus, we examined the role of prior knowledge as a potential moderator or covariate in the analyses of learning outcomes described below.

Did the Conditions Perform Differently on Learning Outcomes?

We conducted a multiple linear regression analysis with condition, prior knowledge, and an interaction term between them as predictors in predicting comprehension and transfer. The ANOVA table showed no significant interaction effect between prior knowledge and condition in predicting comprehension, $F(1,81) = 1.37$, $p = .245$, or transfer, $F(1,81) = .81$, $p = .371$. Therefore, to test the effect of condition on comprehension and transfer, we conducted a univariate analysis of covariance (ANCOVA) with condition as the independent variable and prior knowledge as a covariate in predicting comprehension and transfer. Table 7 presents the adjusted means and standard errors for the learning outcomes. The results indicated a significant difference between conditions in comprehension, $F(1, 82) = 7.08$, $p = .009$, $\eta^2 = .08$, but not in transfer, $F(1, 82) = .68$, $p = .411$. Specifically, the supported drawing condition performed significantly better than the provided visuals condition on the comprehension test, $t(82) = 2.66$, $p = .009$[1], $d = .58$. These findings partially supported our learning outcome hypothesis that drawing with strong support (which combines the benefits of drawing and studying provided visuals) would be *more effective* than only studying the provided visuals.

Did the Conditions Differ in Learning Time and Cognitive Load?

We used two-sample independent t-tests to compare conditions on learning time and cognitive load. Results showed that the visually supported drawing condition spent significantly more time than the provided visuals condition, $t(75.428) = -4.52$, $p < .001$, and the visually supported drawing condition also experienced significantly higher cognitive load than the provided visuals condition, $t(78.24) = -3.22$, $p = .002$.

In sum, these findings supported our learning time and cognitive load hypothesis that the visually supported drawing condition would spend significantly more learning time and experienced significantly higher cognitive load than the provided visuals condition.

Does Learning Time Mediate the Relationship Between Conditions and Learning Outcomes?

To better understand the learning process, we explored why the supported drawing condition performed significantly better than the provided visuals condition on comprehension. Since the supported drawing condition spent significantly more time than the provided visuals

---

[1] Consistent with Experiment 1, drawing quality for the supported drawing condition was significantly positively correlated with comprehension ($r = .66$, $p < .001$) and transfer ($r = .50$, $p = .001$). Furthermore, a paired-samples t-test indicated that students' drawing quality significantly improved after the instructional support, $t(40) = -6.68$, $p < .001$. The means and standard deviations of drawing quality before and after support are shown in Table 5

condition and Table 6 showed that there was a significant positive correlation between learning time and comprehension ($r = .48$, $p < .001$), we tested whether the effect of condition on comprehension is mediated by learning time. We conducted a mediated regression analysis (Figure 1) using bootstrapping, following guidelines from Hayes and Preacher (2014). First, the total effect of conditions on comprehension was statistically significant ($b = 4.42$, $SE = 2.05$, $t(83) = 2.15$, $p = .03$). Next the path from conditions to learning time was statistically significant ($b = 600.34$, $SE = 131.66$, t $(83) = 4.56$, $p < .001$). and the path from learning time to comprehension was statistically significant ($b = .0068$, $SE = .002$, $t(82) = 4.37$, $p < .001$). The path from conditions to comprehension, controlling for learning time, was not statistically significant ($b = .35$, $SE = 2.08$, $t(82) = .17$, $p = .87$). Finally, using a 95% confidence interval obtained from 5000-bootstrap resamples, the analysis indicated the indirect effect of conditions on comprehension through learning time was statistically significant ($b = 4.07$, $SE = 1.36$, 95% *CI* [1.86, 7.11]). This suggests that learning time fully mediated the relationship between the effect of conditions and comprehension.

The mediation analysis indicates that the visually-supported drawing condition performed better on comprehension because they spent more time during learning. It is important to note, as we explain below, the visually-supported drawing condition likely benefited from spending additional time studying the provided visuals as feedback from which to revise their drawings rather than the act of drawing per se (Van Meter, 2001). Therefore, drawing with visual support where students can benefit from *both* drawing activities *and* studying provided visuals as feedback was worth the added time and effort for a better comprehension learning outcome.

Discussion

Empirical Contributions

Prior research has established the multimedia learning effect and the generative drawing effect based on comparisons to students who learn from text or use text-based *without* visualizations. However, little is known about the relative effectiveness of different ways to learn *with* visualizations. Prior research comparing drawing with studying provided visuals only examined effects on learning outcomes without considering their influence on learning time and cognitive load, which is an important consideration both theoretically and practically. The present study addressed this gap by examining learning outcomes as well as learning time and cognitive load. Furthermore, previous studies implemented varying levels of drawing support (i.e., from minimal guidance to comparing and revising one's drawings), likely contributing to the mixed findings (Authors, 2018). The present study compared studying provided visuals to drawing with two forms of high-level support. The primary goal was to determine whether any unique benefits of drawing activities under different conditions were worth the added time and effort.

Experiment 1 provided a high level of *verbal* support (i.e., text-based labels and guided revision questions) without the confounding effect of studying elements from provided visuals. The results partly supported our learning time and cognitive load hypothesis: the verbally-supported drawing condition spent significantly more time and experienced significantly higher cognitive load than the provided visuals condition. However, Experiment 1 did not support our learning outcomes hypothesis. The provided visuals condition performed just as well as the unsupported and verbally-supported drawing condition on both comprehension and transfer. In other words, the provided visuals condition achieved similar learning outcomes with less time and effort. One possible explanation is that the text-based guided revision questions were not sufficient to improve students' judgment accuracy of their drawings and thus, did not help students revise drawings to improve drawing quality.

Experiment 2 implemented a high level of *visual* support for drawing by asking students to compare their drawings to the provided visuals and revise drawings accordingly. The results supported our learning time and cognitive load hypothesis: the visually-supported drawing condition spent significantly more time and experienced significantly higher cognitive load than the provided visuals condition. Experiment 2 also partly supported our learning outcomes hypothesis. Students in the visually-supported drawing condition (who drew *and* studied provided visuals) performed significantly better than students who only studied provided visuals on comprehension, but not on transfer. The mediation analysis showed that the visually-supported drawing condition performed better on comprehension because they spent more time, likely benefiting from the additional time revising their drawings by studying the provided visuals rather than the act of drawing per se. Results from Experiment 1 provide additional support for this explanation. In Experiment 1, in which the verbally-supported drawing condition only received text-based support without opportunities of studying provided visuals, learning time was *not* significantly associated with learning outcomes.

Taken together, the present study adds an important boundary condition to learning by drawing: Whether learning by drawing is worth the added time and effort may depend on different types of drawing support. When students generate drawings *and* have the opportunity to study provided visuals as feedback, the additional time and effort enhances learning. In contrast, drawing activities without visual forms of support may not be worth the added time and effort compared to studying provided visuals.

Theoretical Contributions

Findings from the present study contribute toward integrating theoretical perspectives of traditional research on the multimedia learning effect and the generative drawing effect. Research on these two effects has mostly compared each strategy with learning from text only or text-based strategies (e.g., text vs. text-and-picture, or text vs. text-and-draw). As each strategy involves different cognitive mechanisms, it is important to directly compare different ways of learning *with* visualizations and investigate their unique cognitive benefits and boundaries. The present study helps fill this theoretical gap by comparing the effects of drawing with studying provided visuals on learning outcomes, learning time, and cognitive load. Our results showed that for our sample of students with relatively low prior knowledge of the topic, studying labeled and annotated visuals was generally as effective for learning outcomes as asking students to draw. Although students who drew with visual support (in Experiment 2) outperformed students who studied provided visuals on comprehension, the two groups performed equally well on transfer. This suggests drawing activities—including generating, comparing, and revising one's drawing—may consume too much of students' cognitive resources so that they do not have enough working memory capacity to build a coherent and accurate mental model. In contrast, provided visuals may serve as scaffolds from which students can focus their limited cognitive resources on constructing their mental model.

The findings provide implications for clarifying boundary conditions of learning with visualizations not fully addressed by existing theoretical models. In particular, by testing both verbal and visual forms of drawing support, the results clarify the conditions under which learning by drawing strategy is worth the added time and effort. In Experiment 1, students who drew with pure verbal support spent significantly more time and effort but achieved the same learning outcomes as students who studied the provided visuals. In Experiment 2, students who drew with visual support also spent significantly more time and effort, but the added time yielded higher comprehension performance compared to students who studied provided visuals. Students likely benefited from spending additional time revising drawings by studying the provided visuals rather than drawing per se (Van Meter, 2001). Therefore, drawing activities *alone* may contribute little to learning, especially for novices as in the present study. Only when drawing was combined with visual support such as studying provided visuals as feedback did drawing benefit learning. One possibility is that drawing serves as a preparatory activity that allows students to subsequently process provided visual more productively (Authors, 2019).

Practical Contributions

When comparing different learning or instructional strategies, it is important to consider learning contexts such as student characteristics, features of learning materials, and instructional goals (Ainsworth, 2006; Ainsworth et al., 2016). Accordingly, our findings suggest that if the instructional goal is for low prior knowledge students to learn from complex textual materials, studying labeled and annotated visuals with the text will enable students to achieve similar learning outcomes with less time and effort compared to asking students to draw from the text. Our learning outcome findings suggest that substantial instructional support with visual elements is critical for low prior knowledge students to benefit more from drawing than from studying provided visuals alone. Providing students with visuals as feedback after they finish drawing could help students benefit the most from both strategies and thus, enhance learning.

Limitations and Future Research Directions

One limitation of the study is participants in the two experiments were recruited from different populations. Experiment 1 included students recruited from a mix of introductory-level STEM courses and from an Educational Psychology Subject Pool, whereas all students in

Experiment 2 were recruited from the subject pool. As a result, 39% of students in Experiment 1 were STEM majors, whereas only 4% of students in Experiment 2 were STEM majors. This likely contributed to differences in performance on the prior knowledge test and learning outcomes measures across the two experiments[2]. However, given the large performance differences, we cannot rule out additional factors such as low reliability of the measures across the two experiments, higher interest or engagement among the two samples, or other contextual factors. For those reasons, readers should interpret the present findings with caution, particularly when drawing conclusions that rely on cross-experiment comparisons.

Another apparent limitation of the present study is we did not include a delayed posttest. Learning benefits of the effortful and constructive nature of drawing might be most pronounced on delayed measures of learning compared to studying provided visuals. However, this idea of drawing as a desirable difficulty was not supported in the recent study by Schmidgall and colleagues (2019), in which they found no significant differences between drawing and studying provided visuals on both immediate and delayed transfer tests. Thus, it appears likely that a similar pattern of results would hold on delayed tests, though this should be tested directly in future research.

Moreover, we did not include an unsupported drawing condition as a baseline comparison in Experiment 2 (as we did in Experiment 1), which did not allow us to directly test whether our form of drawing support enhanced drawing quality. We removed the unsupported drawing condition from Experiment 2 because our primary goal was to test an implementation of

[2] We compared individual differences of prior knowledge using the Wilcoxon rank test of median and spatial ability using an independent two-sample t-test between students in two experiments. The results showed that students in Experiment 1 had a significantly higher prior knowledge, $W = 5850$, $p < .001$ and spatial ability scores, $t(166.62) = 2.81$, $p = .005$.

drawing that could “compete” with studying provided visuals. Furthermore, the supported drawing condition in Experiment 2 did outperform the provided illustrations condition on comprehension and students’ drawing quality did improve after guidance, suggesting that the drawing support in Experiment 2 was effective (unlike in Experiment 1). A related criticism might be that we did not include a text-only control condition that did not study or generate visuals. However, as we described above, many prior studies (using similar learning materials) have established that instructor-provided visuals or learner-generated visuals are more effective than only studying text (Mayer, 2014; Van Meter & Firetto, 2013). In contrast, stronger comparisons between effective instructional approaches are relatively rare, yet have greater practical and theoretical significance (Authors, 2018; Chi & Wylie, 2014). In other words, we know visuals are generally more effective than text alone; what we need is a deeper understanding of when and how different types of visuals (such as instructor-provided or learner-generated) lead to more effective learning.

Some may argue that because the information provided to each condition was not equivalent (in Experiment 1), there may have been a ‘teach to the test’ problem in which the provided visual condition had an advantage on the comprehension test. There are several aspects of our design that should assuage this concern. First, all information from the provided visual could be generated directly from what was stated in the text. We were interested in whether it is better to provide students with the corresponding visual information or ask them to spend the extra time and effort to attempt to generate it themselves. Second, the scoring rubrics for the comprehension test were derived entirely from the text, which was identical for all conditions. In other words, students could theoretically receive all possible points on the comprehension test by only reading the text.

Finally, there are two ways to compare effects of learning or instructional strategies on invested time and cognitive load: One is focusing on time and cognitive load *during learning*; the other one is focusing on time and cognitive load *during testing* (Gog & Paas, 2008). The current study only focused on time and cognitive load during learning. However, it is possible that students who drew with more time and effort during learning spent less time and effort during testing. Therefore, time spent and cognitive load during testing would provide additional insight into different effects of strategies. Future research should take into account time and effort invested during both learning and testing.

Although the present study provided high levels of drawing support, it seems that students may need other forms of support to effectively improve drawing quality and the quality of their mental models. Experiment 1 showed that text-based guided revision questions are insufficient for students to effectively revise drawings and, although drawing support in Experiment 2 significantly improved drawing quality, students' drawing quality was still relatively low (mean score of 24.83 out of 46 possible points). Research suggests students with low prior knowledge maintain their incorrect mental models even after repeated exposure to the learning material (Kriz & Hegarty, 2007). This implies that simply showing students the provided visuals as feedback after drawing for comparison and revision may not be sufficient for them to improve drawing quality and revise their mental models. Future research should investigate other ways to facilitate the comparing and revising process such as making specific differences between students' drawings and the provided illustrations salient and explicit (Richland et al., 2017).

Moreover, the effectiveness of an instructional or learning strategy depends on its relevance to achieving a specific goal (Kalyuga & Singh, 2016). The present study showed that when our goal is to help low prior knowledge students learn from complex textual materials, studying labeled and annotated visuals with the text could enable students to achieve similar learning outcomes with less time and cognitive load compared to learning by drawing. However, learning by drawing may be more effective than studying provided visuals for other instructional or learning goals. For example, drawing could be promising for refining mental models after initial learning. Higher prior knowledge and expertise after initial learning will free up students' cognitive capacity to help them benefit from generative activities like drawing (Chen et al., 2015). Moreover, drawing could also be used as a retrieval practice for knowledge consolidation or drawing might serve as a pre-instruction activity (e.g., productive failure, Kapur, 2016) that prepares students to benefit from future learning opportunities. Thus, more research is needed to test how drawing or studying provided visuals, as well as other strategies for learning with visualizations, uniquely contribute to different types of instructional goals.

References

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, *16*(3), 183–198.

Ainsworth, S., Stieff, M., DeSutter, D., Tytler, R., Prain, V., Panagiotopoulos, D., Wigmore, P., Van Joolingen, W., Heijnes, D., Leenaars, F. & Puntambekar, S. (2016). Exploring the value of drawing in learning and assessment, In C. K. Looi et al. (Eds.), *Transforming learning, empowering learners: The International Conference of the Learning Sciences (ICLS) 2016* (Vol. 2; pp. 1082-1089). Singapore: International Society of the Learning Sciences.

Authors (2016)

Authors (2017)

Authors (2018)

Authors (2019)

Butcher, K. R. (2006). Learning from text with diagrams: promoting mental model development and inference generation. *Journal of Educational Psychology*, *98*(1), 182–197.

Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, *14*(1), 5–26.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press.

Castro-Alonso, J. C., & Uttal, D. H. (2018). Spatial ability for university biology education. *International Conference on Applied Human Factors and Ergonomics*, 283–291.

Chen, O., Kalyuga, S., & Sweller, J. (2015). The worked example effect, the generation effect, and element interactivity. *Journal of Educational Psychology*, *107*(3), 689–704.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*(3), 309–319.

de Jong, T., Ainsworth, S., Dobson, M., van der Hulst, A., Levonen, J., Reimann, P., ... & Swaak, J. (1998). Acquiring knowledge in science and mathematics: the use of multiple representations in technology-based learning environments. In M. W. van Someren et al. (Eds.), *Learning with multiple representations* (pp. 9-40). New York: Elsevier.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58.

Glenberg, A. M., & Langston, W. E. (1992). Comprehension of illustrated text: Pictures help to build mental models. *Journal of Memory and Language*, *31*(2), 129–151.

Gog, T. van, & Paas, F. (2008). Instructional efficiency: revisiting the original construct in educational research. *Educational Psychologist*, *43*(1), 16–26.

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.

Hellenbrand, J., Mayer, R. E., Opfermann, M., Schmeck, A., & Leutner, D. (2019). How generative drawing affects the learning process: An eye-tracking analysis. *Applied Cognitive Psychology*, acp.3559.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, *13*(4), 351–371.

Kalyuga, S., & Singh, A.-M. (2016). Rethinking the boundaries of cognitive load theory in complex learning. *Educational Psychology Review*, *28*(4), 831–852.

Kapur, M. (2016). Examining productive failure, productive success, unproductive failure, and unproductive success in learning. *Educational Psychologist*, *51*(2), 289–299.

Kollmer, J., Schleinschok, K., Scheiter, K., & Eitel, A. (2020). Is drawing after learning effective for metacognitive monitoring only when supported by spatial scaffolds? *Instructional Science*, *48*(5), 569–589.

Kriz, S., & Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *International Journal of Human-Computer Studies*, *65*(11), 911–930.

Leopold, C., Doerner, M., Leutner, D., & Dutke, S. (2015). Effects of strategy instructions on learning from text and pictures. *Instructional Science*, *43*(3), 345–364.

Leopold, C., Sumfleth, E., & Leutner, D. (2013). Learning with summaries: Effects of representation mode and type of learning activity on comprehension and transfer. *Learning and Instruction*, *27*, 40–49.

Leutner, D., Leopold, C., & Sumfleth, E. (2009). Cognitive load and science text comprehension: Effects of drawing and mentally imagining text content. *Computers in Human Behavior*, *25*(2), 284–289.

Leutner, D., & Schmeck, A. (2014). The generative drawing principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 433-448). New York: Cambridge University Press.

Marieb, E. N., & Hoehn, K. (2016). *Human anatomy & physiology: Harlow*. Pearson Education Limited.

Mayer, R. E. (1989). Systematic thinking fostered by illustrations in scientific text. *Journal of Educational Psychology*, *81*(2), 240–246.

Mayer, R. E. (Ed.). (2014). *The Cambridge handbook of multimedia learning*: *Second edition*. New York: Cambridge University Press.

Mayer, R. E., Steinhoff, K., Bower, G., & Mars, R. (1995). A generative theory of textbook design: using annotated illustrations to foster meaningful learning of science text. *Educational Technology Research and Development*, *43*(1), 31–43.

Ollerenshaw, A., Aidman, E., & Kidd, G. (1997). Is an illustration always worth ten thousand words? Effects of prior knowledge, learning style and multimedia illustrations on text comprehension. *International Journal of Instructional Media*, *24*(3), 227.

Paas, F. G., & Van Merriënboer, J. J. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors*, *35*(4), 737–743.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, *38*(1), 63–71.

Peeck, J. (1993). Increasing picture effects in learning from illustrated text. *Learning and Instruction*, *3*(3), 227–238.

Rasch, T., & Schnotz, W. (2009). Interactive and non-interactive pictures in multimedia learning environments: Effects on learning outcomes and learning efficiency. *Learning and Instruction*, *19*(5), 411–422.

Reece, J. B., Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., & Jackson, R. B. (2010). *Campbell Biology* (9 edition). Benjamin Cummings.

Renkl, A., & Scheiter, K. (2017). Studying visual displays: how to instructionally support learning. *Educational Psychology Review*, *29*(3), 599–621.

Richland, L. E., Begolli, K. N., Simms, N., Frausel, R. R., & Lyons, E. A. (2017). Supporting mathematical discussions: the roles of comparison and cognitive load. *Educational Psychology Review*, *29*(1), 41–53.

Schmeck, A., Mayer, R. E., Opfermann, M., Pfeiffer, V., & Leutner, D. (2014). Drawing pictures during learning from scientific text: Testing the generative drawing effect and the prognostic drawing effect. *Contemporary Educational Psychology*, *39*(4), 275–286.

Schmidgall, S. P., Eitel, A., & Scheiter, K. (2019). Why do learners who draw perform well? Investigating the role of visualization, generation and externalization in learner-generated drawing. *Learning and Instruction*, *60*, 138-153.

Schnotz, W. (2014). Integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning, 2nd ed* (pp. 72–103). New York: Cambridge University Press.

Schnotz, W., & Kürschner, C. (2008). External and internal representations in the acquisition and use of knowledge: Visualization effects on mental model construction. *Instructional Science*, *36*(3), 175–190.

Schnotz, W., & Wagner, I. (2018). Construction and elaboration of mental models through strategic conjoint processing of text and pictures. *Journal of Educational Psychology*, *110*(6), 850–863.

Schwamborn, A., Thillmann, H., Opfermann, M., & Leutner, D. (2011). Cognitive load and instructionally supported learning with provided and learner-generated visualizations. *Computers in Human Behavior*, *27*(1), 89–93.

Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction*, *13*(2), 227–237.

Seufert, T. (2019). Training for coherence formation when learning from text and picture and the interplay with learners' prior knowledge. *Frontiers in Psychology*, *10*.

Skuballa, I. T., Dammert, A., & Renkl, A. (2018). Two kinds of meaningful multimedia learning: Is cognitive activity alone as good as combined behavioral and cognitive activity? *Learning and Instruction*, *54*, 35–46.

Stern, E., Aprea, C., & Ebner, H. G. (2003). Improving cross-content transfer in text processing by means of active graphical representation. *Learning and Instruction*, *13*(2), 191–203.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, 22, 123–138.

Uttal, D. H., & Cohen, C. A. (2012). Spatial thinking and STEM education: when, why, and how? In *Psychology of Learning and Motivation* (Vol. 57, pp. 147–181). Elsevier.

Van Meter, P. (2001). Drawing construction as a strategy for learning from text. *Journal of Educational Psychology*, *93*(1), 129-140.

Van Meter, P., Aleksic, M., Schwartz, A., & Garner, J. (2006). Learner-generated drawing as a strategy for learning from content area text. *Contemporary Educational Psychology*, *31*(2), 142–166.

Van Meter, P., & Firetto, C. M. (2013). Cognitive model of drawing construction: Learning through the construction of drawings. In G. Schraw et al. (Eds.), *Learning through visual displays* (pp. 247–280). Information Age Publishing.

Van Meter, P., & Garner, J. (2005). The promise and practice of learner-generated drawing: literature review and synthesis. *Educational Psychology Review*, *17*(4), 285–325.

Wiley, J. (2018). Picture this! Effects of photographs, diagrams, animations, and sketching on learning and beliefs about learning from a geoscience text. *Applied Cognitive Psychology*, acp.3495.

Table 1

Means and Standard Deviations Across Groups for Spatial Ability, Prior Knowledge, Cognitive Load, Learning Time, Drawing Quality, and Learning Outcomes for Experiment 1

	Group					
	Provided visuals		Unsupported drawing		Verbally supported drawing	
Measure	*M*	*SD*	*M*	*SD*	*M*	*SD*
Spatial ability (0-10)	6.68	2.03	6.49	2.17	5.97	1.68
Prior knowledge (0-24)	2.97	3.04	3.32	4.08	3.08	4.33
Cognitive load						
Study session 1 (2-14)	9.35	1.47	9.30	1.33	10.31	1.56
Study session 2 (2-14)	8.74	1.75	9.19	1.81	10.11	1.49
Total study time (seconds)	1244.85	471.20	1530.03	536.32	2345.83	1005.12
Drawing quality (0-46)	NA	NA	24.24	10.76	27.19	4.45
Comprehension (0-55)	20.28	10.77	21.00	11.77	18.39	7.65
Transfer (0-29)	5.24	2.93	4.49	2.61	4.92	3.29

Note. The numbers in the parenthesis for each measure represent the range of possible scores for each measure.

Table 2

Pearson Correlations among Individual Difference Variables and Learning Outcome Variables for Experiment 1

	SA	PK	LT	CL	C
Spatial ability					
Prior knowledge	.11				
Learning time	-.14	-.26**			
Cognitive load (averaged over two sessions)	-.31**	-.48**	.31**		
Comprehension	.17	.51**	-.06	-.32**	
Transfer	.23*	.45**	-.04	-.19*	.53**

*$p < .05$; **$p < .01$.

Note. SA = spatial ability; PK = prior knowledge; LT = learning time; CL= cognitive load; C = comprehension.

Table 3

Adjusted Means and Standard Errors across Groups for Comprehension and Transfer Scores for Experiment 1

	Comprehension (0-55)		Transfer (0-29)	
Group	*M*	*SE*	*M*	*SE*
Provided visuals	20.60	1.51	5.21	.45
Unsupported drawing	20.74	1.45	4.39	.43
Verbally supported drawing	18.45	1.47	5.04	.44

Note. Scores adjusted for spatial ability and prior knowledge. The numbers in the parenthesis for each measure represent the range of possible scores for each measure.

Table 4

Means and Standard Errors across Groups for Learning Time and Cognitive Load for Experiment 1

	Learning time (seconds)		Cognitive load (2-14)	
Group	*M*	*SE*	*M*	*SE*
Provided visuals	1245	122	9.04	.24
Unsupported drawing	1530	117	9.24	.23
Verbally supported drawing	2346	119	10.21	.24

Note. The numbers in the parenthesis for each measure represent the range of possible scores for each measure.

Table 5

Means and Standard Deviations Across Groups for Spatial Ability, Prior Knowledge, Cognitive Load, Learning Time, Drawing Quality, and Learning Outcomes for Experiment 2

	Group			
	Provided visuals		Visually supported drawing	
Measure	*M*	*SD*	*M*	*SD*
Spatial ability (0-10)	5.73	2.20	5.24	2.37
Prior knowledge (0-24)	2.00	2.89	1.15	1.53
Cognitive load (2-14)	9.55	1.28	10.54	1.53
Learning time (seconds)	1044.71	528.46	1645.05	680.53
Drawing quality (before support; 0-46)	NA	NA	21.85	8.79
Drawing quality (after support; 0-46)	NA	NA	24.83	8.77
Comprehension (0-55)	11.57	8.12	15.99	10.72
Transfer (0-29)	4.07	3.24	3.29	2.35

Note. The numbers in the parenthesis for each measure represent the range of possible scores for each measure.

Table 6

Pearson Correlations among Individual Difference Variables and Learning Outcome Variables for Experiment 2

	SA	PK	LT	CL	C
Spatial ability					
Prior knowledge	-.12				
Learning time	.03	-.14			
Cognitive load	.05	-.33**	.27*		
Comprehension	.11	.22*	.48**	-.01	
Transfer	.15	.28**	.18	.07	.52**

*$p < .05$; **$p < .01$.

Note. SA = spatial ability; PK = prior knowledge; LT = learning time; C = comprehension; CL= cognitive load.

Table 7

Adjusted Means and Standard Errors across Groups for Comprehension and Transfer Scores for Experiment 2

	Comprehension (0-55)		Transfer (0-29)	
Group	*M*	*SE*	*M*	*SE*
Provided visuals	11.11	1.39	3.94	.42
Visually supported drawing	16.48	1.44	3.43	.44

Note. Scores adjusted for prior knowledge. The numbers in the parenthesis for each measure represent the range of possible scores for each measure.

Table 8

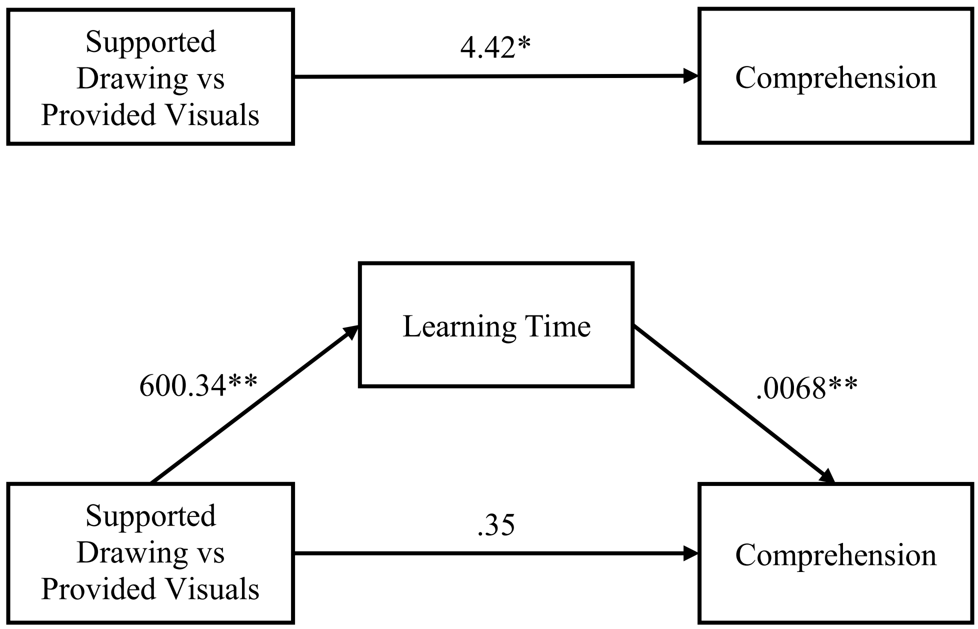
Means and Standard Errors across Groups for Learning Time and Cognitive Load for Experiment 2

	Learning time (seconds)		Cognitive load (2-14)	
Group	*M*	*SE*	*M*	*SE*
Provided visuals	1044.71	132.83	9.55	.31
Visually supported drawing	1645.05	132.83	10.54	.31

Note. The numbers in the parenthesis for each measure represent the range of possible scores for each measure.

Figure 1

Mediation Analysis of Conditions, Learning Time, and Comprehension Test Performance for Experiment 2



*p < .05; **p < .01

Appendix A: Text and provided illustration for human circulatory system lesson

A.1. The Human Circulatory System

The heart has two receiving chambers on the top, the right atrium and left atrium, that receive blood returning from the body and the lungs. The heart also has two main pumping chambers at the bottom, the right ventricle and left ventricle, that pump blood around the circulation.

Although it is a single organ, the heart functions as a double pump. The right side receives relatively oxygen-poor blood from the veins of the upper and lower body through the large superior and inferior venae cavae (ka′ve) respectively and pumps it out through the pulmonary trunk. The pulmonary trunk splits into the right and left pulmonary arteries, which carry blood to the lungs, where oxygen is picked up and carbon dioxide is unloaded. Oxygen-rich blood drains from the lungs and is returned to the left side of the heart through the pulmonary veins. The circulation just described, from the right side of the heart to the lungs and back to the left side of the heart, is called the pulmonary circulation. Its only function is to carry blood to the lungs for gas exchange and then return it to the heart.

Blood returned to the left side of the heart is pumped out of the heart into the aorta, from which the systemic arteries branch to supply essentially all body tissues. Oxygen-poor blood circulates from the tissues back to the right atrium via either the superior or inferior vena cava of the higher or lower body. This second circuit, from the left side of the heart through the body tissues and back to the right side of the heart, is called the systemic circulation. It supplies oxygen- and nutrient-rich blood to all body organs. Because the left ventricle is the systemic pump that pumps blood over a much longer pathway through the body, its walls are substantially thicker than those of the right ventricle, and it is a much more powerful pump.

Why does the heart pump blood in two separate circles? The answer involves two important gases, oxygen (02) and carbon dioxide (C02). All the cells of the body get energy by combining sugars or other food materials with oxygen. This chemical reaction is like burning, which gives off heat and other forms of energy. This energy provides the power we need to talk and move and think. When a body cell combines sugar with oxygen to get energy, carbon dioxide is formed. But too much carbon dioxide could poison a cell. The blood brings oxygen to the body cells and takes away their carbon dioxide. These gases pass easily back and forth through the thin walls of the tiny capillaries in the lungs and the body.

Inside the lungs there are many tiny air sacs called alveoli. They are wrapped in a network of capillaries. As blood flows through the capillaries, the gases on either side of this infinitesimal divide strain to reach equilibrium. On one hand, the pressure of oxygen in the alveolar air is higher than its pressure in the blood, so molecules of oxygen diffuse across the membranes into the blood. On the other hand, the greater pressure of carbon dioxide in the blood forces the gas to flow from the capillaries into the alveoli.

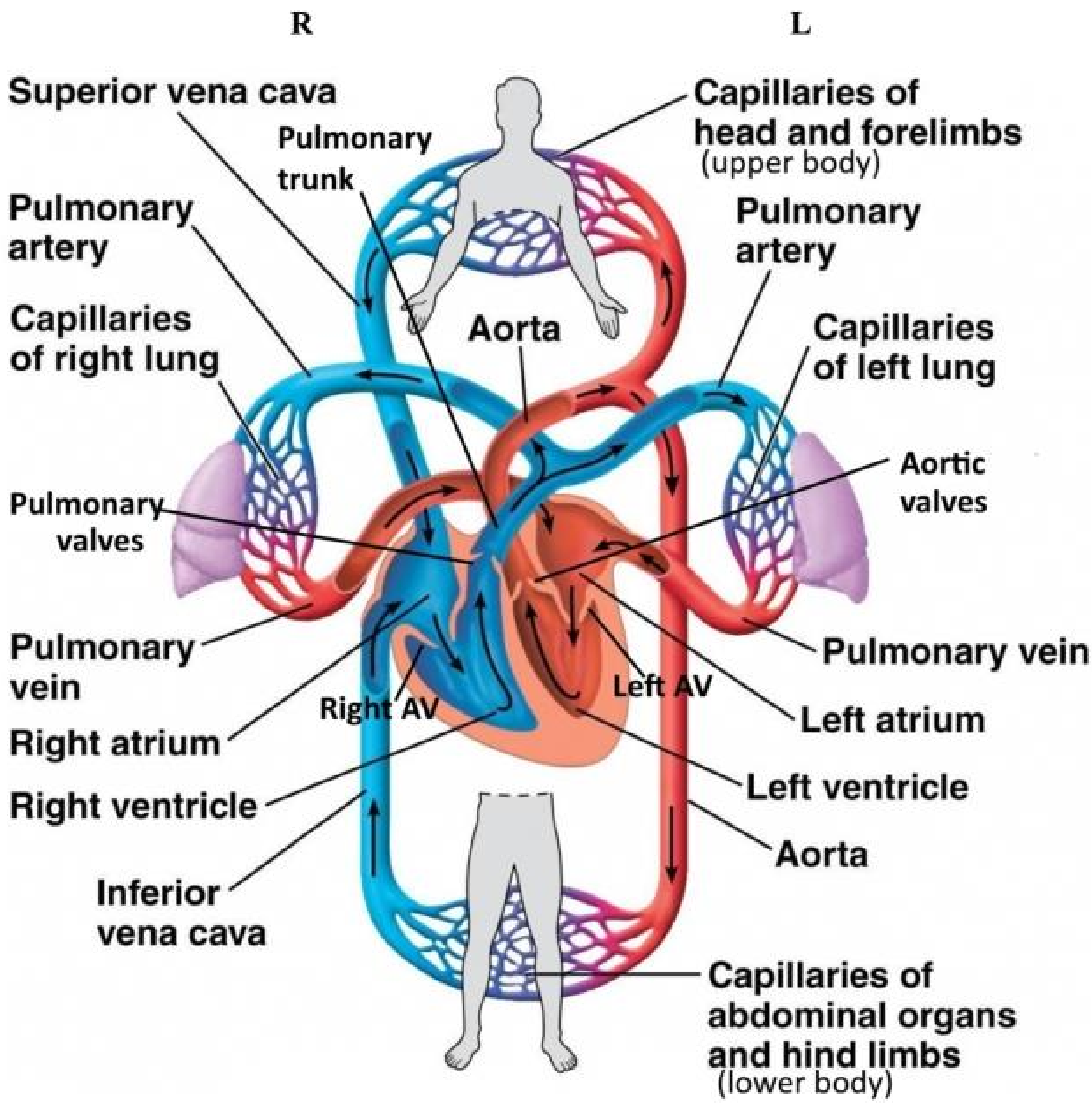
Blood makes a complete circuit from the right side of the heart through the lungs, back through the left side of the heart and out the aorta every two-and-a-half seconds when a person is at rest. During exercise, the blood can travel this short loop in about one second. A remarkable blood protein in the red blood cell, hemoglobin, is largely responsible for the speed with which this vital exchange between oxygen and carbon dioxide occurs. Rich in iron, hemoglobin can unload carbon dioxide and absorb oxygen sixty times faster than blood plasma, the fluid portion of blood. Each molecule of hemoglobin carries four molecules of oxygen to the tissues of the body. It is hemoglobin, when combined with oxygen, that gives blood its bright red color. As the blood is drained of oxygen, the molecules of hemoglobin lose their brilliant red stain, leaving the blood a dull purple, the color it wears on its long journey back through the heart to the lungs.

The heart is equipped with four valves, which allow blood to flow in only one direction through the heart chambers—from the atria through the ventricles and out the arteries leaving the heart. The atrioventricular (a″tre-o-ven-trik′u-lar), or AV, valves are located between the atrial and ventricular chambers on each side. These valves prevent backflow into the atria when the ventricles contract. The left AV valve—the bicuspid, or mitral (mi′tral), valve—consists of two flaps, or cusps. The right AV valve, the tricuspid valve, has three flaps. Tiny white cords, the chordae tendineae (kor′de ten-din′e)—literally, "tendinous cords" (but I like to think of them as the "heart strings" of song)—anchor the flaps to the walls of the ventricles. When the heart is relaxed and blood is passively filling its chambers, the AV-valve flaps hang limply into the ventricles. As the ventricles contract, they press on the blood in their chambers, and the pressure inside the ventricles begins to rise. This forces the AV-valve flaps upward, closing the valves. At this point the chordae tendineae are working to anchor the flaps in a closed position. If the flaps were unanchored, they would blow upward into the atria like an umbrella being turned inside out by a gusty wind. In this manner, the AV valves prevent backflow into the atria when the ventricles are contracting.

The second set of valves, the semilunar (sem″e-lu′nar) valves, guards the bases of the two large arteries leaving the ventricular chambers. Thus, they are known as the pulmonary and aortic valves. Each semilunar valve has three leaflets that fit tightly together when the valves are closed. When the ventricles are contracting and forcing blood out of the heart, the leaflets are forced open and flattened against the walls of the arteries by the tremendous force of rushing blood. Then, when the ventricles relax, the blood begins to flow backward toward the heart, and the leaflets fill with blood, closing the valves. This prevents arterial blood from reentering the heart.

Each set of valves operates at a different time. The AV valves are open during heart relaxation and closed when the ventricles are contracting. The semilunar valves are closed during heart relaxation and are forced open when the ventricles contract. As they open and close in response to pressure changes in the heart, the valves force blood to continually move forward in its journey through the heart.

The Human Circulatory System



Appendix B: Drawing papers in Experiment 1

Drawing paper for the Verbally-Supported Drawing Condition in Experiment 1

Drawing Paper

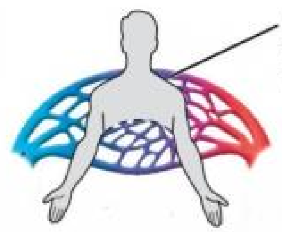
Task: draw a picture by using labels from the label bank to name different parts

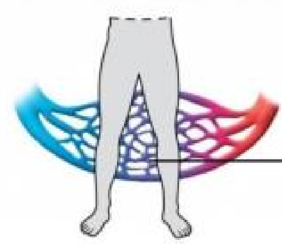
Criteria: 1). use ALL the labels in the label bank; 2). your drawing should correctly reflect relations between different parts described in the lesson so that others could understand how the human circulatory system works from your drawings.

Use lines with arrows (→) to represent the direction of blood flow in the vessels and the heart
Use lines (——) to label different parts

R **L**

The Human Circulatory System





Label Bank

left atrium **right atrium** **left ventricle** **right ventricle** **left lung** **right lung**	**superior venae cavae** **inferior venae cavae** **pulmonary trunk** **pulmonary arteries** **pulmonary veins** **aorta**	**pulmonary valves** **aortic valves** **left AV(or bicuspid/mitral valve)** **right AV (or tricuspid valve)**

Drawing paper for the Unsupported Drawing Condition in Experiment 1

Drawing Paper

Draw a picture or multiple pictures on the drawing paper to help you understand how the circulatory system works

R **L**

Appendix C: Drawing paper for the Visually-Support Drawing Condition in Experiment 2

Drawing Paper

Criteria: 1). include all essential parts described in the text; 2). your drawing should correctly reflect relations between different parts described in the lesson so that others could understand how the human circulatory system works from your drawings.

Use lines with arrows (⟶) to represent the direction of blood flow in the vessels and the heart
Use lines (——) to label different parts

R **L**

Appendix D: Instructions in Experiment 1

Study Session 1:

The provided visuals condition:
Please carefully study the lesson explaining the human circulatory system. Later you will be tested on what you learned. Please take as long as you need and let me know when you are done. Do you have any questions? Do you understand the task?

The unsupported drawing condition:
Please carefully study the lesson explaining the human circulatory system. Later you will be tested on what you learned. To help you learn from the lesson, draw a picture or multiple pictures on the drawing paper to help you understand how the circulatory system works. Please take as long as you need and let me know when you are done. Do you have any questions? Do you understand the task?

The verbally-supported drawing condition:
Please carefully study the lesson explaining the human circulatory system. Later you will be tested on what you learned. To help you learn from the lesson, draw a picture by using labels from the label bank to name different parts. Your completed drawings should meet two criteria: (1) you should use all the labels in the label bank; and (2) your completed drawing should correctly reflect relations between different parts described in the lesson so that others could understand how human circulatory system works from your drawings. Please take as long as you need and let me know when you are done. Do you have any questions? Do you understand the task?

Study Session 2:

The provided visuals condition:
Please study the lesson again. Then you will be tested on what you learned. Please take as long as you need and let me know when you are done.

The unsupported drawing condition:
Please study the lesson again; you can revise your drawings if necessary. Then you will be tested on what you learned. Please take as long as you need and let me know when you are done.

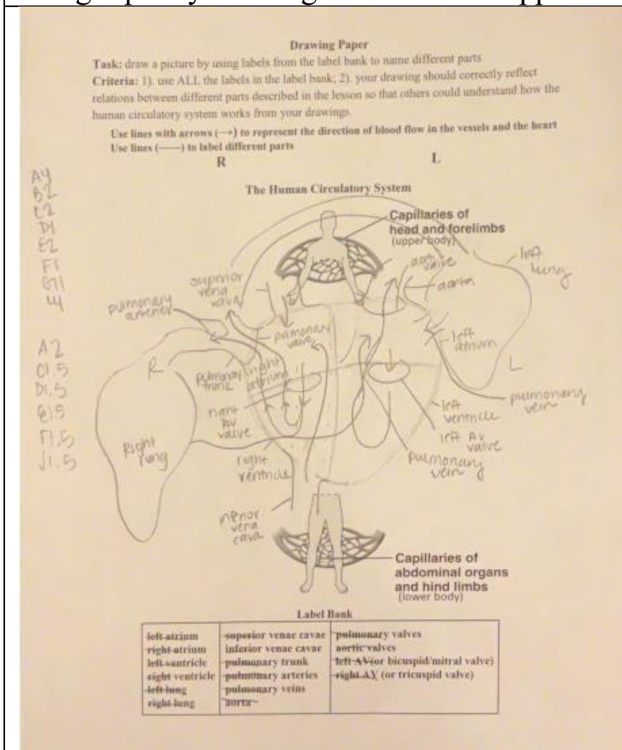
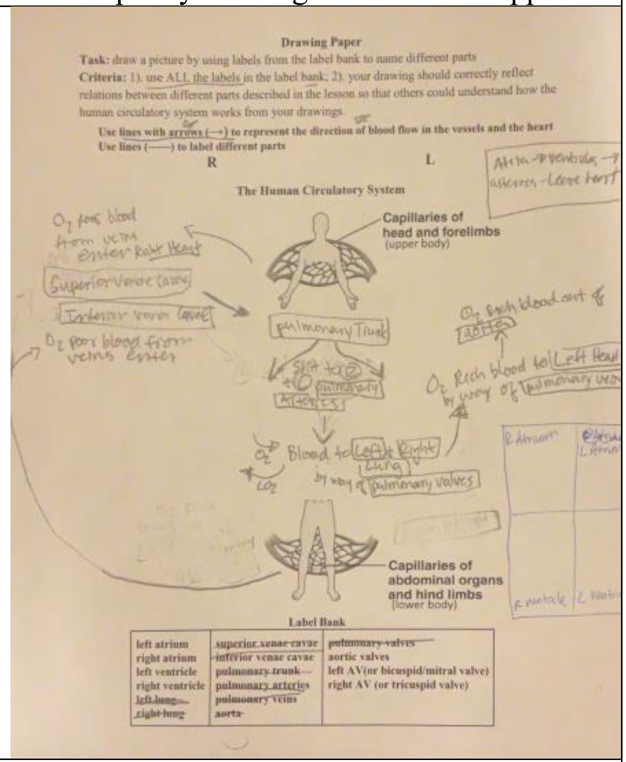
The verbally supported drawing condition:
Please check your drawings again by answering the following questions. Based on your responses to the questions, revise your drawings as necessary. Then you will be tested on what you learned. Please take as long as you need and let me know when you are done.

Appendix E: Instructions in Experiment 2

The provided visuals condition:
Please carefully study the lesson explaining the human circulatory system. Later you will be tested on what you learned. Please take as long as you need and let me know when you are done. Do you have any questions? Do you understand the task?

The visually supported drawing condition:
Please carefully study the lesson explaining the human circulatory system. Later you will be tested on what you learned. To help you learn from the lesson, draw a picture or multiple pictures on the drawing paper. Your completed drawings should meet two criteria: 1) you should include all essential parts described in the text; 2) your completed picture should correctly reflect relations between different parts described in the lesson so that others could understand how human circulatory system works from your drawings. Please take as long as you need and let me know when you are done. Do you have any questions? Do you understand the task?
[After providing feedback] Please check your completed drawings by comparing to the provided picture. Revise your drawings as necessary. Then you will be tested on what you learned. Please take as long as you need and let me know when you are done

Appendix F: Examples of Students' Drawings in Experiment 1

High quality drawings with verbal support	Low quality drawings with verbal support
	
High quality drawings without support	Low quality drawings without support