

# UC Berkeley

## UC Berkeley Electronic Theses and Dissertations

### Title

Independent Study Strategies for Learning about the Cardiovascular System from Text: A Comparison of Self-explanation and Drawing

### Permalink

<https://escholarship.org/uc/item/7351f5bs>

### Author

Lam, Diane Phuong N

### Publication Date

2015

Peer reviewed|Thesis/dissertation

Independent Study Strategies for Learning about the Cardiovascular System from Text:  
A Comparison of Self-explanation and Drawing

by

Diane Phuong Nghinh Lam

A dissertation submitted in partial satisfaction of the

Requirements for the degree of

Doctor in Philosophy

in

Science and Mathematics Education

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Angelica M. Stacy, Chair

Professor Andrea A. diSessa

Professor William C. Sha

Summer 2015

Copyright 2015 Diane Lam

## Abstract

### Independent Study Strategies for Learning about the Cardiovascular System from Text: A Comparison of Self-explanation and Drawing

by

Diane Phuong Nghinh Lam

Doctor of Philosophy in Science and Mathematics Education

University of California, Berkeley

Professor Angelica Stacy, Chair

Representations, such as figures and drawings, are aspects of biology that are key to learning, teaching, and communicating scientific ideas. While many studies have investigated undergraduate students' abilities to interpret representations generated by science experts, much remains to be understood about how student-generated representations (i.e., drawings) can support learning. Prior research suggests that theoretical mechanisms to explain how drawing aids learning may parallel those that explain how self-explanation aids learning, an area of educational research that has been extensively studied. As such, this research draws from the self-explanation literature to explore the similarities and differences between the use of drawing and self-explaining as independent study strategies for learning about the cardiovascular system (CVS) from text.

We found that students who were asked to draw as they studied the CVS text performed better than students asked to self-explain on multiple learning measures. Their mental models, as interpreted from their drawings of the system, were significantly more accurate, and their responses to questions about structures and pathways within the system were more accurate. Further analyses of self-explanations and drawings revealed that the number of goal-oriented self-explanations a student generated was a significant predictor of their assessment scores, especially on questions about functions. Meanwhile, the number of passage-specific images a learner generated in their drawing was predictive of assessment scores, especially on questions about structures. Finally, findings from case studies identified attributes of self-explanations and drawings that may make them more meaningful for learning, such as self-explaining for the purposes of understanding how parts of the system connect together, and drawing to highlight the main ideas of the text. Findings from this study suggest that drawing is generally more effective than self-explanation for learning about the CVS, and should be promoted as a useful strategy for learning biology from text.

*This dissertation is dedicated to the memory of Randi A. Engle*

## **ACKNOWLEDGEMENTS**

---

I would never have been able to complete this project without the help of my amazing undergraduate research assistants. Thank you Kristina Yim, Echo Lu, Elizabeth Sabiniano, Michael DeChenne, Norielle Adricula, Alex Tseng, Tammy Yabiku, Brandon Duong, and Anusha Challa. You all helped in so many different ways and really made the process fun. I would also like to thank my advisors Angy Stacy, Andy diSessa, and Bill Sha for all of your guidance and support through all parts of my entire program. I thank my classmates who supported me in and out of school, especially Adi Adiredja, Sarah Menanix, and Jen Wang. Finally, thank you Leslea Hlusko, not only for letting me collect data in your classroom but also for all of your advice and kindness.

# TABLE OF CONTENTS

---

<b>1</b>	<b>INTRODUCTION AND OVERVIEW .....</b>	<b>1</b>
<b>2</b>	<b>BACKGROUND.....</b>	<b>2</b>
2.1	<b>What is a drawing? .....</b>	<b>2</b>
2.2	<b>Interpreting and Generating Inscriptions .....</b>	<b>2</b>
2.2.1	The use of inscriptions in biology courses .....	2
2.2.2	Interpreting versus generating inscriptions .....	3
2.3	<b>Generating Inscriptions and Generating Self-Explanations.....</b>	<b>5</b>
2.3.1	What is self-explaining?.....	5
2.3.2	Relationships between self-explaining and drawing.....	6
2.4	<b>Active and Constructive Learning.....</b>	<b>7</b>
2.5	<b>Research Questions.....</b>	<b>7</b>
<b>3</b>	<b>Methods .....</b>	<b>8</b>
3.1	<b>Subject Domain .....</b>	<b>8</b>
3.2	<b>Participants.....</b>	<b>9</b>
3.3	<b>Experimental Conditions .....</b>	<b>10</b>
3.4	<b>Procedure.....</b>	<b>11</b>
3.4.1	Recruitment and Screening.....	12
3.4.2	Pre-Assessment .....	13
3.4.3	Assigning Conditions.....	13
3.4.4	Day 1 .....	14
3.4.5	Day 2.....	15
3.5	<b>Materials .....</b>	<b>15</b>
3.5.1	Screening Survey .....	15
3.5.2	Training Materials .....	15
3.5.3	CVS Text Booklets .....	16
3.5.4	Formal Assessments .....	17
3.6	<b>Data Sources and Reorganization.....</b>	<b>19</b>
3.6.1	Formal assessments .....	19
3.6.2	Independent study sessions .....	20
<b>4</b>	<b>Part 1: Formal Assessments.....</b>	<b>22</b>
4.1	<b>Drawing Task.....</b>	<b>22</b>
4.1.1	Analyses of Drawing Mental Models from Formal Assessment Drawings.....	22
4.1.2	Formal Assessment Drawing Results.....	26
4.2	<b>Free-response Questions.....</b>	<b>29</b>
4.2.1	Analysis of Free-Response Questions.....	29
4.2.2	Free-Response Results .....	33
4.3	<b>Figure Interpretation Task.....</b>	<b>40</b>
4.3.1	Labels .....	41
4.3.2	The Direction of Blood Flow.....	42
4.4	<b>Time on Task .....</b>	<b>46</b>
4.5	<b>Discussion of Formal Assessment Results .....</b>	<b>48</b>
4.5.1	Rewriting .....	49
4.5.2	Self-explaining .....	50
4.5.3	Drawing.....	53

4.5.4	Combined – Self-Explaining and Drawing .....	58
<b>5</b>	<b>Part 2: Use of Study Strategies .....</b>	<b>60</b>
<b>5.1</b>	<b>Analyses of Self-Explanations .....</b>	<b>60</b>
5.1.1	Grain size.....	60
5.1.2	Defining self-explanations .....	61
5.1.3	Examples of self-explanations.....	62
<b>5.2</b>	<b>Results of Self-Explanation Analyses.....</b>	<b>64</b>
5.2.1	Pooled: Self-explanations from the SE and Combined groups.....	65
5.2.2	Separately: Self-explanations from the SE and Combined groups .....	71
<b>5.3</b>	<b>Analyses of Booklet Drawings .....</b>	<b>74</b>
<b>5.4</b>	<b>Results of Booklet Drawings Analyses .....</b>	<b>76</b>
5.4.1	Pooled: Drawings from the Draw and Combined groups.....	77
5.4.2	Separately: Drawings from the Draw and Combined groups .....	79
<b>5.5</b>	<b>Discussion of Self-explanation and Drawing Results.....</b>	<b>81</b>
5.5.1	Self-explaining .....	81
5.5.2	Drawing.....	85
5.5.3	Self-explaining vs. Drawing .....	87
<b>6</b>	<b>Part 3: Case Studies .....</b>	<b>89</b>
<b>6.1</b>	<b>Analyses of Self-explanations and Drawings .....</b>	<b>89</b>
6.1.1	Quality of Self-explanations .....	89
6.1.2	Quality of Drawings.....	90
<b>6.2</b>	<b>Identifying Participants for Case Studies .....</b>	<b>91</b>
<b>6.3</b>	<b>Adam: High Learning Gains.....</b>	<b>92</b>
6.3.1	Adam’s background.....	93
6.3.2	Adam before studying (Pre-test results) .....	93
6.3.3	Adam after studying (Post and delayed test results) .....	94
6.3.4	Adam’s opinions of the study strategies .....	97
6.3.5	Adam’s Self-explanations.....	98
6.3.6	Adam’s Drawings .....	103
<b>6.4</b>	<b>Beth: Average Learning Gains .....</b>	<b>110</b>
6.4.1	Beth before studying (Pre-test results).....	111
6.4.2	Beth after studying (Post and delayed test results).....	112
6.4.3	Beth’s opinions of the study strategies.....	116
6.4.4	Beth’s Self-explanations .....	116
6.4.5	Beth’s Drawings .....	122
<b>6.5</b>	<b>Discussion .....</b>	<b>128</b>
6.5.1	Personal backgrounds.....	129
6.5.2	Attention to aesthetics.....	129
6.5.3	Self-explaining .....	130
6.5.4	Drawing .....	132
6.5.5	Other comments .....	134
<b>7</b>	<b>Overall discussion.....</b>	<b>136</b>
<b>7.1</b>	<b>Summary of Findings .....</b>	<b>136</b>
7.1.1	Part 1: Formal Assessments.....	136
7.1.2	Part 2: Use of Study Strategies .....	136
7.1.3	Part 3: Case Studies.....	136



7.2	<b>A Return to the Research Questions</b> .....	137
7.2.1	Research Question #1.....	137
7.2.2	Research Question #2.....	137
7.2.3	Research Question #3.....	138
7.2.4	Research Question #4.....	139
7.2.5	Research Question #5.....	139
7.3	<b>Limitations</b> .....	139
8	<b>Conclusion</b> .....	141
9	<b>References</b> .....	142
10	<b>Appendix</b> .....	147
10.1	<b>Appendix A: Screening Survey</b> .....	147
10.2	<b>Appendix B: Training PowerPoint Slides</b> .....	151
10.3	<b>Appendix C: Cardiovascular System Text</b> .....	155
10.4	<b>Appendix D: Free-Response Questions</b> .....	157
10.5	<b>Appendix E: Pre-test MANOVA Assumptions</b> .....	159
10.6	<b>Appendix F: Post-test QSA ANCOVA Assumptions</b> .....	160
10.7	<b>Appendix G: Delayed-test QSA ANCOVA Assumptions</b> .....	162
10.8	<b>Appendix H: Post-test QSB ANCOVA Assumptions</b> .....	164
10.9	<b>Appendix I: Delayed-test QSB ANCOVA Assumptions</b> .....	166
10.10	<b>Appendix J: Correlation Matrix of Time on Task and Sub-scores</b> .....	168
10.11	<b>Appendix K: Correlation Analyses of Full Self-explanation Data</b> .....	170

## LIST OF FIGURES

Figure 1: Hierarchy of terms describing representations.	2
Figure 2: Order of events for participants in the study.	12
Figure 3: Figure presented to participants for the figure interpretation task.	19
Figure 4: The six mental models of the cardiovascular system identified by Chi et al (1994). Figure adapted from Friedman & Forbus (2011).	23
Figure 5: Example drawing that does not indicate that the lungs play a role in the oxygenation of blood (#0491 Post-test)	24
Figure 6: Mean drawing mental model pre-test scores with standard errors.	27
Figure 7: Mean short-term drawing mental model gain scores with standard errors.	28
Figure 8: Mean long-term mental model drawing gain scores with standard errors.	28
Figure 9: A. Mean pre-test scores (QSA) with standard errors. B. Mean pre-test (QSA) sub- scores by structure, pathway, and function with standard errors.	36
Figure 10: A. Adjusted mean post-test QSA scores with standard errors. B. Adjusted mean post-test QSA SPF sub-scores with standard errors. Asterisks indicate significant differences between groups, determined by post hoc t-tests.	37
Figure 11: A. Adjusted mean delayed-test QSA scores with standard errors. B. Adjusted mean delayed-test QSA SPF sub-scores with standard errors.	38
Figure 12: A. Adjusted mean post-test QSA scores with standard errors. B. Adjusted mean post-test QSB SPF sub-scores with standard errors.	38
Figure 13: Adjusted mean delayed-test QSB scores with standard errors. B. Adjusted mean delayed-test QSB SPF sub-scores with standard errors.	39
Figure 14: Combined bar graphs for SPF sub-scores for the post-test (Figures 4B and 6B) and delayed-test (Figures 5B and 7B), by question set. These means were adjusted to control for pre-test sub-scores.	40
Figure 15: Mean figure interpretation task labeling scores with standard errors	42
Figure 16: Ten positions indicating the correct direction of blood flow through the CVS.	43
Figure 17: Mean blood flow scores with standard errors.	44
Figure 18: Mean times spent studying the material in total and perceptibly using a study strategy, and standard errors. All post hoc comparisons between pairs of groups were significant at the $\alpha = 0.01$ level except where indicated as not significant (NS).	46
Figure 19: Combined bar graphs for SPF sub-scores for the post-test and delayed-test, by question set. These means (and standard errors) were adjusted to control for pre-test sub-scores and time spent using study strategies.	48
Figure 20: Box and whiskers plots for A. total number of self-explanations and B. self- explanation formats (goal-driven, elaborative, and noticing coherence). Outliers circled in red were not included in the correlation analyses.	66
Figure 21: Scatterplots with regression lines for A. structures sub-scores and B. functions sub-scores, predicted by the number of goal-driven self-explanations generated.	68
Figure 22: Box and whiskers plots for mentions of structures, pathways, and functions in self-explanations. Outliers circled in red were dropped from correlation analyses.	69
Figure 23: Scatterplots with regression lines of A. structures sub-scores, B. pathways sub- scores, and C. functions sub-scores, predicted by the number of times functions are mentioned in self-explanations.	71
Figure 24: Coding scheme items for drawings generated while studying Passages 3 and 11. Each drawing was evaluated for the presence of a label or image representing each of	

these items. Items marked with (I) were evaluated for the presence of an image only, and items marked with (L) were evaluated for the presence of a label only.	75
Figure 25: The three arrows point to examples of labels generated by participant #9830 while studying the CVS text.	76
Figure 26: Box and whiskers plots of labels and images from drawing data pooled from the Draw and the Combined groups	77
Figure 27: Scatterplots and regression lines for A. structures sub-scores, B. functions sub-scores, and C. DMM scores predicted by the number of images generated.	79
Figure 28: Adam and Beth had similar pre-test scores and total times spent studying.	92
Figure 29: While Adam and Beth had similar pre-test scores, they had different overall post and delayed test scores.	92
Figure 30: Adam's pre-test drawing of the CVS.	94
Figure 31: Adam's post and delayed test drawings of the CVS.	96
Figure 32: Adam's drawing of Passage 2, the general structure of the heart.	104
Figure 33: Participant #1648's drawing of Passage 8, representative of most participants drawing of this passage.	105
Figure 34: Adam's drawing of Passage 8, the process of diffusion through a membrane.	106
Figure 35: Adam's drawing of Passage 6, the structure of blood vessels.	107
Figure 36: Adam's drawing of Passage 9, the function of valves.	108
Figure 37: Adam's drawing of Passage 10, the two primary subsystems of circulation.	109
Figure 38: Adam's drawing of Passage 4, phases of muscle contraction.	110
Figure 39: Beth's pre-test drawing of the CVS.	112
Figure 40: Beth's drawings of the CVS from the post and delayed tests	115
Figure 41: Beth's drawing of Passage 2, the general structure of the heart.	124
Figure 42: Beth's drawing of Passage 8, the process of diffusion through a membrane.	125
Figure 43: Beth's drawing of Passage 9, the function of valves.	126
Figure 44: Beth's drawing of Passage 10, the two primary subsystems of circulation.	127
Figure 45: Beth's drawing of Passage 4, phases of muscle contraction.	128

## LIST OF TABLES

Table 1: Participant demographics .....	10
Table 2: Types and number of questions on each formal assessment based on Chi et al. (1994) .....	18
Table 3: Various times measured from video data.....	21
Table 4: Mental model coding scheme for drawings on formal assessments. ....	24
Table 5: Data summary of drawing mental model scores (and standard deviation) by assessment, based on a 6-point scale. ....	27
Table 6: Number of opportunities for participants to display SPF knowledge on QSA and QSB free-response questions.....	33
Table 7: Data summary of unadjusted mean scores (and standard deviations) for each group on each formal assessment. Scores are each presented as a percentage of the total possible points for that overall score or sub-score. Sub-scores: S = Structures, P = Pathways, F = Functions. ....	35
Table 8: Labels for figure interpretation task that recieved full or partial credit.....	41
Table 9: Mean scores (and standard deviations) for labeling components on figure interpretation task.....	42
Table 10: Paths of blood flow categorized as correct or by type of error. ....	45
Table 11: Summary of blood flow direction indicated through position 4 of participants with incorrectly drawn blood paths. ....	45
Table 12: Correlation matrix between times using study strategies and formal assessment scores.....	47
Table 13: Coding of self-explanations (with standard deviations) for participants in the SE group and the Combined group .....	65
Table 14: Correlation matrix of formal assessment sub-scores and DMM scores versus total self-explanations and self-explanation formats for the pooled data from the SE and Combined groups. ....	67
Table 15: Correlation matrix of formal assessment sub-scores and DMM scores versus mentions of SPF categories in self-explanations for the SE and Combined groups. ....	69
Table 16: Correlation matrices of formal assessment sub-scores and DMM scores versus self-explanation formats from A. the SE group and B. the Combined group.....	72
Table 17: Correlation matrices of formal assessment sub-scores and DMM scores versus mentions of SPF categories in self-explanations from participants in A. the SE group and B. the Combined group. ....	73
Table 18: Coding of drawings (with standard deviations) for participants in the Draw group and the Combined group.....	77
Table 19: Correlation matrix of formal assessment sub-scores and DMM scores versus drawing labels and images for the pooled data from the Draw and Combined groups. 78	
Table 20: Correlation matrices of formal assessment sub-scores and DMM scores versus the number of labels and images generated by participants in the A. Draw group, and B. Combined group. ....	80
Table 21: Adam’s pre-test scores, along with the average scores for all participants and those only in the Combined group. ....	93
Table 22: Adam’s averaged post and delayed test scores, and average scores for all participants and those only in the Combined group.....	95
Table 23: Beth and Adam’s pre-test scores, along with the average scores for all participants and those only in the Combined group. ....	111

**Table 24: Beth and Adam’s average post and delated test scores compared to average scores for all participants and those only in the Combined group.....113**  
**Table 25: Beth’s and Adam’s booklet drawing scores compared to average scores for all participants and those only in the Combined group..... 122**

# 1 INTRODUCTION AND OVERVIEW

---

Whether or not they decide to become research biologists, undergraduate students majoring in biology will undoubtedly need to make decisions involving complex biology as citizens at various times in their lives. Even those who do not choose to study biology will still need to make decisions at the doctor's office or at the voting booth, and these decisions will require some degree of content knowledge and an understanding of how biology is communicated (AAAS, 2011; Kloser, 2010). Thus, considering that millions of college students are likely to, at the very least, take a biology course as part of their general education requirements, they need to understand how to learn in them.

With developments in graphic technology over the last several decades, biology, now more than ever, has become largely embodied by the inscriptions (graphs, diagrams, illustrations, etc.) that are essential to the scientific, educational, and everyday worlds that use them (O'Donoghue et al., 2010; Schönborn & Anderson, 2006). These inscriptions not only help people to learn and think about scientific ideas (Cox & Brna, 1995; Kindfield, 1994), but also to communicate and argue ideas with each other (Kozma, 2003; Latour, 1990; Roth & McGinn, 1998). It would help students learning biology to recognize the importance of these inscriptions in their studies.

Also, considering the many options available for taking classes today, students would benefit from study strategies that help them to learn independently. Online classes are becoming ubiquitous and many universities offer online webcasts of lectures, resulting in many students learning without the physical presence of an instructor and classmates. Even in a typical college course where students go with their classmates to a physical classroom multiple times a week to listen to the instructor lecture, a significant portion of learning still occurs outside of the classroom when students are studying with their notes and textbooks. It would help students studying biology to be aware of effective study techniques for learning from these notes and textbooks.

The focus of this dissertation is to consider ways in which undergraduate students can learn about a biology topic from text. Specifically, we will investigate how two independent study strategies may support student learning: self-explaining and drawing. Part 1 of this dissertation takes a largely quantitative approach to studying the effects of these strategies by comparing learning (based on formal assessment scores) across groups of students trained to use the strategies. Part 2 takes a more detailed look at the self-explanations and drawings that students generated while studying by categorizing formats of each and finding relationships between them and learning. Finally, Part 3 presents two case studies of students using both study strategies and describes qualities of the self-explanations and drawings they generated that were likely to be useful for learning.

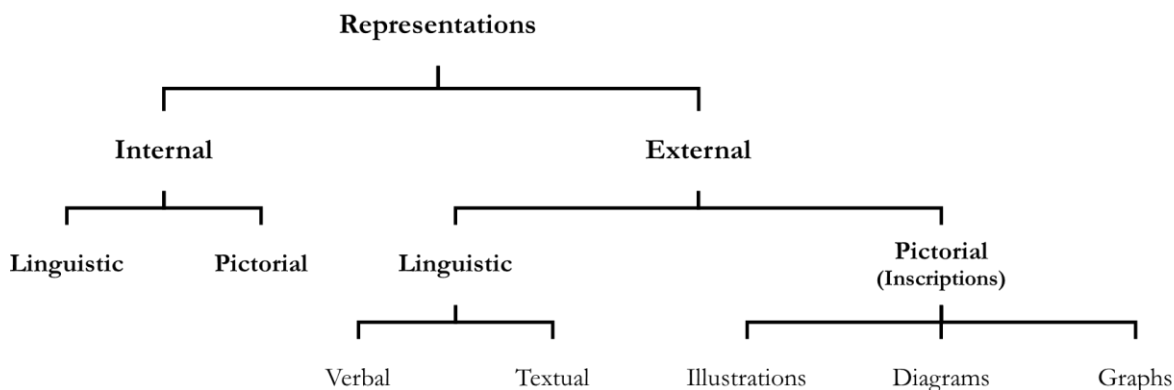
## 2 BACKGROUND

---

### 2.1 What is a drawing?

---

For the purposes of this dissertation, it is important that we first consider meanings of several related terms that have been used in the literature surrounding representations. While there are inconsistencies in how these terms have been used across the literature, they will be used in the following ways throughout this dissertation. A visual hierarchy of the terms are presented below in Figure 1. Most broadly, the term “representation” will refer to both the internal and external forms of information. In other words, it refers to both how information is imagined in one’s mind and how information is presented in the world, respectively (Scaife & Rogers, 1996). More specifically, “inscriptions” refer to external pictorial representations such as illustrations, diagrams, and graphs, and excludes external representations of linguistic forms (Roth & McGinn, 1998).



**Figure 1:** Hierarchy of terms describing representations.

Inscriptions may be generated on a computer or by hand. In this study, participants were asked to generate their inscriptions by hand, with pen and paper. We will simply refer to these as drawings. This research places a spotlight on hand-drawn inscriptions but expands on the empirical and theoretical studies of these other levels of representations to understand how student learning may be impacted by representations, in general.

### 2.2 Interpreting and Generating Inscriptions

---

#### 2.2.1 The use of inscriptions in biology courses

---

Undergraduate biology students are typically positioned as passive learners in the classroom (Dirks, 2011; Kliensky, 2001; Knight & Wood, 2005). It is also reasonable to assume that most college biology courses lack explicit instruction to develop student abilities to fully engage with scientific inscriptions (Schönborn & Anderson, 2009). Opportunities for learning with inscriptions do exist in these courses, such as exposure to an abundance of inscriptions in course textbooks and instruction involving inscriptions during class. However, these opportunities are usually poorly structured, in terms of supporting learning, and do not usually reflect the inscriptions and practices seen in authentic science (Kindfield & Singer-Gabella, 2010).

More often than not, textbooks are required purchases for students in college biology courses, and are filled with inscriptions that accompany the text. However, the mere presence of these inscriptions is not sufficient for getting students to attend to them, let alone try to learn from them. Students admit to ignoring the figures presented in the margins of their textbooks, even when they are specifically referenced in the text (Linn & Eylon, 2011). To make matters worse, the inscriptions used in these textbooks are not representative of those that appear in authentic science. While inauthentic inscriptions may be better for instructional purposes, exposing students to those seen in the science field can, to some extent, prepare them for interacting with these types of inscriptions in the future. Textbooks tend to use photographs and diagrams to accompany the text, whereas science journals tend to use more mathematical figures such as tables and graphs (Bowen & Roth, 2002; Roth, Bowen, & McGinn, 1999). Thus, not only are inscriptions in textbooks generally overlooked by students, but they are not representative of the types of inscriptions that they would need to interpret in the science world or will come across in the media.

With the guidance of an instructor, development of inscriptional practices may not be much better supported than by studying independently from a textbook. Kindfield & Singer-Gabella (2010) describe classroom observations of biology instructors modeling poor inscriptional practices during instruction. Rather than using and presenting the inscriptions as tools of practice, instructors in the study presented the inscriptions as transparent facts, meant to transmit information. Thus, it is not surprising that many students who have graduated from college have been shown to incorrectly or inadequately interpret graphs in biology, even when their undergraduate degrees were conferred in biology (Bowen, Roth, & McGinn, 1999).

In addition to skills interpreting inscriptions, other representational skills have been identified in the literature that appear relatively unsupported in undergraduate biology education. These include, but are not limited to, being able to translate between different forms of inscriptions (Kozma & Russell, 2005; Stieff, 2011), critiquing inscriptions based on suitability for tasks (diSessa, 2004), and being metacognitive about the information presented in external representations (Gilbert, 2005). Unfortunately, these skills are largely left uncultivated, leaving a wealth of learning opportunities about inscriptions untouched or only vaguely presented (Schönborn & Anderson, 2006). As a result, many students simply glance over figures in their textbooks, lecture notes, and journal articles and move on, underestimating the opportunities presented to them to learn and the demands required to engage in learning from inscriptions (Linn & Eylon, 2011).

### 2.2.2 Interpreting versus generating inscriptions

---

As we consider these other skills surrounding the ability to interpret inscriptions, we cannot avoid the fact that interpretation is only the “reading” component to understanding the language of biology inscriptions. In order to develop true representational fluency or literacy, one must also learn how to “write,” or generate inscriptions. In biology, the generation of inscriptions grounds scientific practice as well as scientific knowledge, and can serve both individual and social functions (Latour, 1990). For an individual, inscriptions can act as tools for cognitive support, allowing for greater computational efficiency, re-representation, and graphical constraining (Larkin & Simon, 1987). Studies have shown that chemists and biologists use and create inscriptions to assist in their personal thinking about scientific phenomena (Kindfield, 1994; Kozma, 2003; Stieff, Hegarty, & Deslongchamps, 2011). Additionally, in social contexts, inscriptions have been shown to support collaboration and argumentation, as people share



information or present claims about science (Kozma, 2003; Latour, 1990; Roth & McGinn, 1998).

Research on learner-generated inscriptions (i.e., drawing) has followed the surge in research on student interpretation of inscriptions. However, much of how drawing impacts learning remains elusive. diSessa, Hammer, Sherin, & Kolpakowski (1991) demonstrated that young students have remarkable abilities to generate inscriptions of scientific ideas. Assigned the task of representing motion, sixth grade students (almost independent of their teacher) collaborated to iteratively refine a classroom inscription that described a motorist speeding up, stopping and then driving away. While it was remarkable that their final version of an inscription to represent motion developed into a speed versus time graph, what was more interesting was the range of competence that these young students displayed about representations when discussing their ideas. Not only were they able to generate unique inscriptions, but they considered aspects of design such as advantages, disadvantages, consistency, flexibility, variety, and purpose. While this study showed, over two decades ago, that the act of generating an inscription in this context provided students with opportunities to learn simultaneously about both representations and motion, not enough studies have expanded on this idea.

Martin & Schwartz (2009) showed that graduate and undergraduate students (not medically trained) managed representations differently when tasked with making medical diagnoses from large amounts of data. The graduate students were found willing to spend the extra time to make prospective adaptations to their representations in anticipation of future benefits, while undergraduate students tended to make adaptations reactively when as they encountered problems. Within the framework of novice to expert understanding, the decisions that the graduate students made were representative of adaptive expertise (Hatano & Inagaki, 1986). What makes the findings from this study particularly interesting is that participants in neither group of this study were experts in the medical diagnosing. It suggests that the adaptive expertise exhibited by the graduate students was more specific to representations, and less to the content. The graduate students had more experience managing and analyzing data than the undergraduate students and were thus better able to plan their representations to be useful to them in problem contexts. We can imagine that having adaptive expertise in representing would be beneficial for becoming an expert in any domain that uses representations.

In another study, analyses of “workscratching” on a GRE-like test revealed great diversity among the inscriptions that students generated (Cox & Brna, 1995). They found that the best construction method was based on the utility of drawings at the “read-off stage.” In other words, the phase that drawings were found to be most useful for problem-solving was after they were fully constructed and were referred to for answering the question. This is in contrast to the phase before drawing begins, when the learning is preparing for what he or she will draw, and the phase of the actual generation of the inscription. In the current study, students are not asked to solve specific problems with their drawings, per se, but understanding the content material is a problem in itself. Thus, the read-off stage may be particularly important for our participants as well.

Research on how drawing can aid learning for *undergraduate biology students* is particularly sparse. One study explored how new and senior biology undergraduates students compared to expert geneticists in their spontaneous use of drawings to solve genetics problems (Kindfield, 1994). The more advanced participants in this study designed their inscriptions of DNA specifically for each problem context and used them to support their computations, record information, check previous reasoning, and to illustrate components of their solution.

Meanwhile, the less advanced students tended to include extraneous details in their drawings, and even when the student included all of the features necessary to answer the question, their inscriptions did not adequately support their reasoning to arrive at the correct solution. The study suggested that conceptual understanding and inscriptional skills coevolve. That is, the more a learner understands the material, the better his or her inscriptional practices become, and vice versa.

## 2.3 Generating Inscriptions and Generating Self-Explanations

---

Alongside the empirical research investigating when generating inscriptions aids learning, researchers are also considering theoretical explanations for *how* it aids learning. One explanation that many researchers are considering relates the process of drawing-to-learn with the process of self-explaining (Ainsworth & Loizou, 2003; Cox & Brna, 1995; Cox, 1997; Cromley, Newcombe, & Bergey, 2011).

### 2.3.1 What is self-explaining?

---

#### *a. Defining self-explaining*

Self-explaining is the act of explaining learning material to oneself. It is specifically the act of talking to oneself and not to another person, and making inferences about the material one is studying. Many previous studies have attempted to operationalize self-explanations but have done so with some variation.

The main criteria that distinguishes a self-explanation from any other utterance is that it is inferential. That is to say, it must present information that is not explicitly stated in the learning material. For example, when studying material from a text, a student can generate a self-explanation in three different ways: 1) by integrating information in the text with his or her prior knowledge 2) by integrating information across different sentences in the text, and 3) by using the meaning of words in the text to infer what may also be true.

Some prior studies on self-explanation have also considered metacognitive monitoring statements to be self-explanations (e.g., Ainsworth & Loizou, 2003; Cromley et al., 2011; Roy & Chi, 2005). However, the original self-explanation study (Chi, Bassok, Lewis, Reimann, & Glaser, 1989) and a follow-up study focused on learning about the cardiovascular system (Chi, Leeuw, Chiu, & LaVancher, 1994) did not. In this study, we do not consider monitoring statements to be self-explanations because they are generally not inferences about the material. More details about coding self-explanations will be provided in later sections, but here we will investigate what self-explanations do for learning.

#### *b. Generating inferences and revising mental models*

The self-explanation effect was first described by Chi et al. (1989). Through the analysis of think-aloud protocols, the main finding was that “Good students,” who performed better than “Poor students” on problem solving tasks, tended to generate more self-explanations as they studied worked-out examples of mechanics problems. This effect has been replicated in the context of studying worked-out probability calculations (Renkl, 1997), sections of a geometry textbook (Wong, Lawson, & Keeves, 2002), computer programming lessons (Bielaczyc, Pirolli, & Brown, 1995), and text about the cardiovascular system (Chi, 1994).

Two mechanisms have been proposed to explain the self-explanation effect: through generating inferences, and through revising mental models (Chi et al., 1989; Chi et al., 1994;

Chi, 2000). The generation of inferences is generally a more overt activity, and was described in the previous section. The more complicated and often covert activity is revising mental models.

Gentner & Stevens (1983) describe a mental model as an internal representation of an inter-connected system of concepts that relates to the external structure it represents (as cited in Chi, 2008). Chi (2008) also describes mental models as “[a]n organized collection of individual beliefs” (p.67). By these descriptions, a mental model is similar to a coordination class, as defined by diSessa (2002). Each describes a system or network of connected fragments that can constitute a model of a scientific concept. Both mental models and coordination classes have also been described to have incorrect versions, where they can be misaligned with the correct scientific model.

In this dissertation, we will use “mental model” to describe this network of information because Chi (2008) places a heavy emphasis on the ability to run mental models, “much like an animated simulation,” which is significant for the system our participants learn about (p. 67). Furthermore, coordination classes have been described to contain mental models and sometimes be adaptable and transferable to other contexts (Amin, Smith, & Wisner, 2014). In this study, student mental models are specific to the one biological system.

### 2.3.2 Relationships between self-explaining and drawing

---

Generating inferences and revising mental models are surely involved in the process of generating an inscription, as well as with self-explaining. Specifically, when drawing from information presented in text, a learner will often need to make inferences about the relative sizes and shapes of components, their spatial relationships to each other, and other details not made explicit in the text. Furthermore, a drawing may be a physical extension of the learner’s mental model. While a physical drawing will not contain all of the fragments of knowledge that a learner has about and related to the system, it can contain some important pieces of it. Thus, as learners add to and edit their physical drawings, they are likely to be simultaneously editing their internal mental models.

Furthermore, consider the three main theoretical points for how self-explanation supports learning (Chi, 1994; Roy & Chi, 2005; VanLehn, Jones, & Chi, 1992):

1. It is a constructive process: In contrast to simply re-reading or repeating the information aloud, generating a self-explanation requires learners to actively form connections, predictions, hypotheses, etc.
2. It promotes integration of existing knowledge and new information: In order to make sense of or contextualize new information, learners are likely to build their explanations from existing knowledge. This is in contrast to activities such as summarizing, which may not necessarily require any connection to other knowledge.
3. It is a continuous and piecemeal process: While studying, individual pieces of new information can be incorporated into existing or developing mental models. Every explanation (and partial explanation) generated by the incorporation of a piece of new information presents an opportunity to see conflicts within the learner’s mental model and correct them.

We can see how the process of drawing to learn can parallel the process of self-explaining to learn. Drawing 1) is a physically and mentally constructive process, 2) involves integrating new information into a developing inscription, and 3) provides learners with opportunities to see conflict in their model, with each addition to their drawing, (cf. Van Meter, Aleksic, Schwartz, & Garner, 2006).

## 2.4 Active and Constructive Learning

---

Let us focus for a moment on the fact that both self-explanation and drawing are overtly constructive processes for learning, and both are expected to be better for learning than any that only qualify as active or passive (Chi, 2009). In this study, some participants are asked to self-explain or draw to learn from text. Others are asked to rewrite the text (that is, hand-copy the words onto a piece of paper), and the rest are asked to both self-explain and draw. In comparison to self-explanation or drawing, rewriting is less of a constructive process and considered to be only active. It is indeed physically constructive but not necessarily mentally constructive. Thus, rewriting is hypothesized to be less effective for learning than self-explaining or drawing. If we ignore fatigue, using both the self-explanation and drawing strategies is believed to be the most constructive and, therefore, the most effective for learning.

Between self-explanation and drawing, I hypothesize that drawing will be more constructive because a learner asked to draw, can easily also self-explain covertly. However, if a student is asked to self-explain, he or she cannot also draw. Furthermore, prior research has shown that students tend to generate more self-explanations from diagrams than from text (Ainsworth & Loizou, 2003). Thus, participants who draw may be inclined to internally self-explain. Essentially, those who are asked to draw may, to some extent, be using both strategies, and may perform similarly to participants explicitly asked to do both. This is considering the fact that participants are asked to learn from text, requiring them to process the textual information to some extent in order to generate a self-explanation or drawing from it. This processing may include some form of internal self-explaining, thus resulting in students who are asked to draw to actually do both. Overall, rewriting is presumed to be an active process, self-explaining is constructive, drawing is more constructive, and self-explaining and drawing together is most constructive.

## 2.5 Research Questions

---

This study aims to investigate what aspects of drawing support student learning and how they compare to self-explanation as a strategy for learning from text. The learning material used in this study consists of passages of text describing the cardiovascular system. The specific research questions are:

- 1) As independent study strategies for overall learning about the cardiovascular system from text, how effective is drawing in comparison to self-explaining?
- 2) Do students who draw learn about different aspects about the cardiovascular system than students who self-explain? For example, do they learn better about more visual aspects of the system?
- 3) Does self-explaining and drawing together have additive or synergistic effects?
- 4) What types of activities do students engage in when they are drawing to learn and do any of those activities correlate to learning gains? In other words, while “Good students” tend to self-explain more when they are asked to think aloud (Chi et al., 1989), do they tend to do something analogous in their drawings?
- 5) What specific attributes of self-explanations and drawings make them meaningful for learning?

### 3 METHODS

---

This section describes the details of the study conducted. It includes an explanation for the choice of biology topic, a general description of the participants, and a detailed description of the procedures. The section concludes with a summary of the types of data that were collected and how data were organized for analysis.

#### 3.1 Subject Domain

---

Complex systems involve relationships and interactions between individual elements that give rise to a collective behavior. Making sense of such systems is often a difficult process for learners because many elements within a system may seem disparate and may function at different organizational levels (Kloser, 2010). Furthermore, forming a unifying framework for the whole system requires an understanding of how each element contributes to the system, sometimes simultaneously, which potentially places a heavy load on one's working memory (Hmelo, Holton, & Kolodner, 2000). However, developing an understanding of complex systems is becoming increasingly important for students as conceptual coherence is central to science learning (Jacobson & Wilensky, 2006).

While complex systems are pervasive in the world in general, an understanding of these types of systems is an important skill in biology, specifically, where organization and emergent processes may be considered unique to this particular domain. As explained in Mayr (2004) about the uniqueness of biology,

Many properties of systems, such as higher levels of integration, cannot be explained by a study of their isolated components. The integration of systems results in the emergence of new properties because 'the whole is [often] more than the sum of the parts.' The emergence of new properties is characteristic of higher levels in any hierarchy of systems (p.102)

Given these complexities, investigating how students learn about biological systems provides rich insight into the challenges that students face and aspects of their learning that can be enhanced with interventions.

The human cardiovascular system (CVS) is a central, complex system in biology and is typically a topic covered in general high school biology courses. The system involves an intricate network of organs and tissues that work together to maintain the circulation of blood and supply nutrients to every cell in the body. A complete understanding of how the human CVS works requires one to recognize each of the components that contribute to the system, how each contributes to the system individually, and how each interacts with one another to generate a collective outcome. Furthermore, one must understand that the dysfunction of an individual component or an interaction between components will have both local and systemic effects. It is a complex and highly connected system.

Prior research has described common erroneous mental models that participants have generated while learning about the human CVS by reading text. Some learners only interpret one loop in the system when two circulatory loops actually exist (one between the heart and lungs, and another between the heart and the rest of the body). Additionally, while some learners may understand how individual organs or tissues play a role in their local parts of the system, some fail to understand how the components contribute to the system as a whole or how they interact with other components (Ainsworth & Loizou, 2003; Chi, 1994).

This study builds upon prior research that has investigated student understanding of the CVS after generating self-explanations from reading text (Chi, 1994), and after reading text or interpreting figures (Ainsworth & Loizou, 2003). The same CVS text was used in these prior studies were used in this study.

## 3.2 Participants

---

All participants were recruited from the same lecture course at an R1 University in the Fall semester of 2013. This course focused on evolutionary biology from an anthropological standpoint (focusing a bit more on sociocultural factors than biological ones) and did not include any material about the CVS. The course fulfilled an American Cultures course requirement for many undergraduate students, and had more variation in students' major, year in school, and age (Table 1) than typical courses that fulfill major requirements. The majority of participants in this study were female (68%), underclassmen (65%), 18 or 19 years of age (65%), and Asian or Pacific Islander (51%). The plurality of participants were social science majors (40%). Of the 288 students in the class, 97 who were invited to participate completed all components of the study.

Data from one of the 97 participants was dropped from the study as an outlier because her scores and sub-scores were the lowest of all participants on each assessment. Review of video data revealed that this participant spent the least amount of time in her study session and on each of the assessments. I believe this participant rushed through the sessions and answering the assessment questions. In the end, data from 96 participants were used for analyses.

**Table 1:** Participant demographics

<u>Demographic Factor</u>	<u>n</u>
Gender	
Female	65
Male	31
Major	
Arts and Humanities	4
Biological Sciences	16
Business	5
Mathematics, Engineering, and Technology	16
Physical Science	7
Social Science	39
Double Major	3
Undeclared	6
Year	
Freshman	32
Sophomore	31
Junior	20
Senior	13
Age	
18	31
19	32
20	19
21+	14
Ethnicity or Race	
Black or African American	1
Asian or Pacific Islander	49
Hispanic or Latino	12
White	27
Multiple Selected	6
Declined to state	1

### 3.3 Experimental Conditions

---

Participants were assigned to one of four conditions in this study: 1) Rewrite 2) Self-Explain (SE), 3) Draw, and 4) Self-Explain and then Draw (Combined). The Rewrite group was instructed to copy the CVS text verbatim, with pen and paper, as they read through the material. The SE group was trained to self-explain, the Drawing group was trained to draw-to-learn, and the Combined group received a combination of both training protocols. I believe having participants rewrite the CVS text reflects a study strategy that students actually use in school, such as when they copy notes that the instructor writes on the board in class. This study strategy allowed students to spend a substantial amount of time with the material and was easily observable in videos.

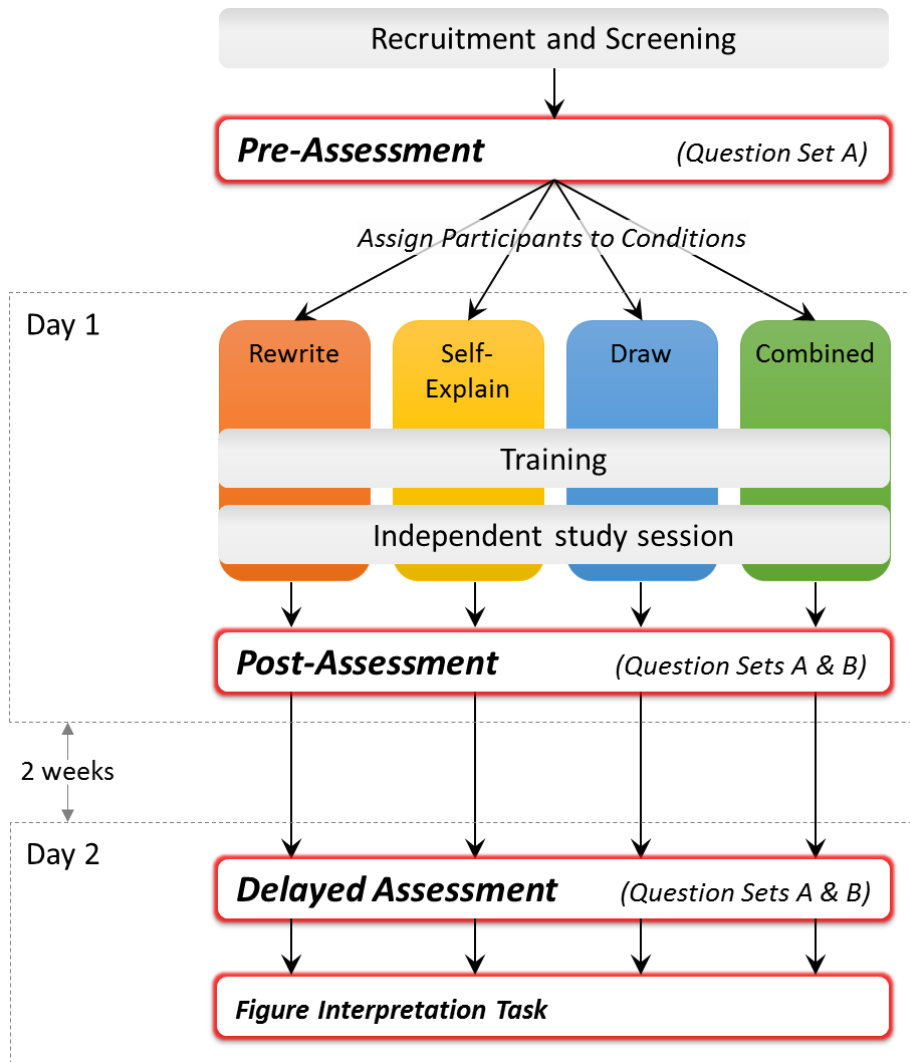
The inclusion of a Rewrite group was originally meant to serve as a negative control group, motivated by the control group used in Chi, Leeuw, Chiu, & LaVancher (1994). In their study, Chi and colleagues had participants first read the CVS text aloud and then either self-explain aloud or re-read the text aloud. The reasons for not repeating the conditions for this control group were twofold. First, it was shown in their study that participants in the control group spent much less time on task than participants who were asked to self-explain. In an attempt to have groups spend similar amounts of time on task, I chose to have participants complete a task that would allow them to spend more time with the CVS text than re-reading aloud. Second, I wanted to provide participants with the option to read the text silently, which many preferred. If they were to read and re-read silently in this study, it would be difficult to monitor their utilization of the study strategy. Rewriting the text was an overt task that achieved both of these aims.

### **3.4 Procedure**

---

This section describes the procedure of the study and briefly introduces the formal assessments that were used to measure learning. Further details about assessments and other materials will be described later in section 3.5. Figure 2 provides a visual overview of the chronological order of events. Each event is described in detail in the following sections.





**Figure 2:** Order of events for participants in the study.

### 3.4.1 Recruitment and Screening

Arrangements were made with the professor of the course from which students were recruited to offer up to 5 points of extra credit (less than 1% of a total grade for the course) for participating in this research study. During the second week of the Fall 2013 semester, at the beginning of a course lecture, I introduced the class to the researchers they would later meet if they participated in the study (myself and two undergraduate research assistants), and provided a brief summary of the study.

The study was presented as a research project focused on study strategies that students can use while learning from text. The specific study strategies (i.e., self-explanation and drawing) were not mentioned, but an emphasis was made about our interest in strategies that could help students learn while reading textbooks. All students in the class were then told to expect a follow-up email that would repeat these details about the study and include additional instructions on how to participate if they chose to.

Following this verbal recruitment, I sent an email to all students enrolled in the class (through the course website) with the same details I presented in lecture, and a link to an online screening survey. Students were asked if they were interested in participating in the study to

complete the screening survey in order for our research team to determine their eligibility. Details of the screening survey are described in section 3.5.1 below. The main purpose of the survey was to identify students who were unlikely to have studied the CVS in the past year and would benefit the most from participating in the study. In other words, students who reported taking a biology class within the last year that likely covered the CVS as a topic (e.g., AP biology, physiology, or anatomy) were not invited to participate. Based on this criterion, 120 students who completed the screening survey were invited to schedule a date to complete the pre-assessment (pre-test).

Students who were not eligible or did not want to participate in the study were offered an alternate (and point-equivalent) extra credit opportunity to read a biology educational research article and write a one-page paper responding to a prompt about the article.

### 3.4.2 Pre-Assessment

---

Students who were invited to participate in the study were emailed several dates and times, and asked to select a session to complete the pre-test. The pre-test was comprised of a set of questions which will be referred to as Question Set A (QSA). At each of these sessions, up to 15 participants met with me in a classroom and sat at desks along the perimeter of the room, facing the walls. Students were each given a blank sheet of paper and three colored pens, and asked to access the online pre-test on individual laptops. They were told to make educated guesses if they were unsure of answers and to use “I don’t know” as a response only if they needed to. They were also asked to remain seated until they were notified that all participants had completed their assessments. This and having students face the walls was meant to deter them from watching others’ progress and prevent those who took more time from feeling rushed. Details of the pre-test (QSA) are provided in section 3.5.4.

After participants completed their assessments, they were asked not to research the answers to the questions or to study the CVS before the next scheduled meeting, which would be scheduled individually via email. Participants were informed that they would need to schedule two subsequent one-on-one meetings for approximately two hours each, two weeks apart. One hundred and two students completed the pre-test.

### 3.4.3 Assigning Conditions

---

Based on a rough scoring of the pre-test responses, participants were ordered in a list according to their total scores and grouped into quartets. The four participants with the highest scores on the list were considered one group, the next four with the next highest scores were considered another group, and so on. The website [www.random.org](http://www.random.org) was then used to generate a random order of the four experimental conditions: Rewrite, SE, Draw, and Combined. Then, four participants in a group would be assigned to a condition based on that random order. A new order of conditions was randomly generated for each group on the list.

This method was used to assign participants to conditions so that the average pre-test scores across the conditions would be similar. Essentially, each experimental condition consisted of one of the top four performers, one of the bottom four performers, and one of every group of four in between. Participants were unaware of the groupings and unaware that multiple types of learning strategies were being used with different participants until the end of the study.

While the pre-tests were administered in a group setting, the remainder of the study was conducted one-on-one. Upon arrival to the Day 1 session, each participant was asked if he or she had learned anything about the CVS since the pre-test. They were assured that there was no penalty if they had and that the fact would simply need to be recorded. No participants reported learning about the CVS between completing their pre-test and their Day 1 sessions. Next, the experimenter summarized that the Day 1 session would consist of a brief training on how to use a study strategy, an independent study session where the participant would use the strategy, and would conclude with an assessment.

#### *a. Training*

Participants in the Rewrite group received a short scripted talk to motivate the use of rewriting as a technique to study, mentioning how students often copy what instructors write on the board in class. In the other three conditions, the experimenter used a PowerPoint presentation to train each participant in using their assigned study strategy. These presentations a) highlighted key features of the strategy; b) included guiding questions to help remind the participant how to use the strategy; c) exhibited an example of a student using the strategy; and d) provided a sample text for the participant to read and practice using the strategy. Detailed descriptions of the training materials are presented in section 3.5.1. After the training presentation was complete, the participant was given an opportunity to ask questions. Participants in the SE, Draw, and Combined groups were provided with a paper printout of the guiding questions from their respective training presentations which they could refer to during their independent study sessions.

#### *b. Independent study session*

Each participant was then provided with a booklet that consisted of a cover page of instructions and 11 passages of text describing the CVS (see Appendix C: Cardiovascular System Text). The experimenter read the instructions to the participant aloud, which explained that the text would be presented passage by passage on each page in the booklet, and that the student was to read and study each passage using the strategy learned from training. The experimenter remained seated with the participant until he finished reading and using the study strategy with the first passage. A cardboard divider was then set up and the participant was left to study the remainder of the booklet (10 passages) alone.

Once finished with the booklet, the participant was asked for opinions about the study strategy, and was requested to provide both positive and negative feedback about the strategy. The participant was then offered a break.

#### *c. Post-assessment*

Once the participant was ready to continue, he or she was provided with a computer with the online post-test cued, a piece of paper, and three colored pens. The participant was asked to attempt to answer all of the questions on the assessment and use the response “I don’t know” only if needed. The post-test was comprised of the same set of questions that appeared on the pre-test (QSA), and also a new set of questions that the student had not previously seen in the study, Question Set B (QSB). These questions did not appear as two consecutive sets, but were mixed together. When the assessment was completed, the participant was reminded of his Day 2 meeting scheduled two weeks from that day. Participants were again asked to not research any of

the answers to the assessment questions or any information about the CVS, in general, for those two weeks.

#### 3.4.5 Day 2

---

Upon returning for the Day 2 session, each participant was asked if he or she had learned anything about the CVS since the end of the Day 1 session. They were assured that there was no penalty if they had and that the fact would simply need to be recorded. No participants reported learning about the CVS between their Day 1 and Day 2 sessions.

##### *a. Delayed assessment and figure interpretation task*

The participant was then provided with a computer with the online delayed assessment cued, a piece of paper, and three colored pens. He or she was asked to attempt to answer all of the questions on the assessment and use the response “I don’t know” only if needed. The delayed test was comprised of QSA and QSB with an additional set of figure interpretation questions never presented before. These final questions involved a diagram of the CVS and asked participants to identify the components indicated and draw the direction of blood flow throughout the system.

##### *b. Debrief*

When finished with the assessment, the participant was asked if he or she had any questions about the CVS or about the research project. If the student was interested, the experimenter would describe the four conditions and overall goals of the study.

### 3.5 Materials

---

This section provides detailed descriptions of the materials used in this study. It includes materials that were used for data collection as well as for the training sessions.

#### 3.5.1 Screening Survey

---

The screening survey included multiple choice, open-response, and Likert-scale questions. The first set of questions asked participants to report to what extent they agreed that they were visual, auditory, reading/writing, and kinesthetic learners. The next set asked students about how much experience they had reading science text. Then, students were asked about the most recent biology course they had completed. Finally, questions about general demographic information were placed at the end of the survey to reduce stereotype threat (Steele & Aronson, 1995). The complete screening survey is provided in [Appendix A](#).

#### 3.5.2 Training Materials

---

The training sessions for each condition were scripted for the experimenter to read and, except for the Rewrite group, were accompanied with PowerPoint presentations. Screenshots of the training slides are provided in [Appendix B](#). The training protocols for each of the conditions were as follows.

##### *a. Rewrite group*

The “training” for this condition did not involve actual practice writing. Instead, the experimenter told the participant that research has shown that copying text can help with memory and comprehension. The script also related the rewriting strategy to what often happens

during course lectures when students copy what an instructor writes on the board. Participants were asked to focus on this rewriting strategy as they read through the CVS booklet and learned the material.

*b. Self-explain, Draw, and Combined groups*

The PowerPoint-guided training sessions for the SE, Draw and Combined groups followed parallel structures. Participants were told that their study strategy has been shown in educational research to enhance problem-solving and comprehension. They were also told that using the strategy involved more than simply repeating the information presented in the text and required making inferences about the material. Participants were instructed to use the study strategy for themselves (not for the experimenter), and to not worry about the eloquence of their self-explanations or the aesthetics of their drawings.

The experimenter then read aloud three guiding questions meant to support self-explaining and/or drawing. Guiding questions for self-explaining pointed to understanding the main idea of the passage and making inferences, relating to other ideas in the text or to prior knowledge, and asking and answering questions about the material. Guiding questions for drawing pointed to representing a coherent picture with complete and connected parts, relating to other ideas in the text or to prior knowledge, and drawing ideas together to form a functional model of the system. The term “model” was elaborated by the experimenter as a useful tool to simulate what is happening or predict what can happen. Participants in the Combined group were presented with both sets of guiding questions.

The next slide in each training PowerPoint presented a table of descriptions of what was considered a self-explanation and what was not, a table of what was considered aspects to a good drawing or not, or both tables. The experimenter read all of these descriptions to the participant. During pilot studies, participants tended to become silent if they were told to avoid rewording or paraphrasing the text, as those are not considered self-explanations. Thus, a note was made for experimenters to explain to participants in the SE and Combined groups that rewording and paraphrasing are not considered self-explanations but were not discouraged because they may sometimes lead to self-explanations.

Finally, participants were presented with two example texts. The first was presented alongside an example of a student using the study strategy, and the second was presented for the participant to practice using the strategy him- or herself. The first example text was a short passage about how penguins maintain their internal body temperature and included either a transcript of a student’s self-explanation, a copy of student’s drawing, or both. After allowing the participant to look over the student example(s), the experimenter followed the script to ask the participant to critique all parts of the self-explanation and/or drawing, adding comments if the participant did not discuss key features. The second example text was a short passage about how camels have features that allow them to survive in the desert. The experimenter asked the participant to read the text, and then practice generating their own self-explanation, drawing, or both. Participants were reminded to use the guiding questions for assistance if they wanted to. When finished practicing, the experimenter made positive comments about their use of the strategy and suggested ways they could improve when necessary.

### 3.5.3 CVS Text Booklets

---

After training was complete and participants asked any questions they had, they were presented with a CVS text booklet. On the front of each booklet were instructions that the

experimenter read aloud and explained the following. Each left page in the booklet contained one passage of text about the CVS. The participant was to read each passage while using the study strategy learned during training. They were instructed to not refer back to previous pages as they proceeded through the booklet to encourage a testing effect, or better long-term retention by providing opportunities for them to retrieve information as they studied (Chan & McDermott, 2007). Participants were also reminded to self-explain and/or draw for themselves and to not worry about eloquence or aesthetics. Participants were also asked to study the material as if they were studying for an exam for class. The instructions then explained that the experimenter would set up a cardboard divider and would wear headphones to give the participant some privacy while he or she studied. The instructions conclude with a prompt for the participant to ask any questions.

Once all of the participant's questions were answered, the experimenter demonstrated how the booklet was organized so that the back of the cover page contained the first passage of the text. This page was to be turned and placed to the left of the remaining pages in the booklet. On the top of each right-hand page was a sentence that asked the participant to copy the text, SE, draw, or SE and then draw. Under that sentence in booklets for the Rewrite group were lines that the student could copy the text onto. In booklets for the other three groups, that space was empty and was either used for drawing or remained unused (SE group). At the bottom of each right-hand page was a sentence telling the student to proceed to the next passage once they had finished using their study strategy to learn from that passage, except for the last page that told them to notify the experimenter that they were finished. A full copy of the CVS text is provided in [Appendix C](#).

#### 3.5.4 Formal Assessments

---

The pre, post, and delayed tests were administered through Google Forms. The first item that appeared on each of the assessments was the drawing task. Participants were provided with three colored pens and a blank piece of paper, and asked to draw the path of blood flow from the heart to the other parts of the body. For the remainder of the questions, except for the final question on the delayed assessment, participants were asked to type their free-response answers into text boxes. The final question on the delayed assessment asked students to draw the path of blood flow onto a print-out figure of the CVS.

##### *a. Question sets*

QSA contained 18 free-response questions and appeared on the pre, post, and delayed tests. QSB contained 16 questions and only appeared on the post and delayed tests. The total 34 questions were chosen from the 43 used in Ainsworth & Loizou (2003), which were chosen from the 80 originally developed and used by Chi, Leeuw, Chiu, & LaVancher (1994). These questions were selected with an effort to cover the most material with the least amount of repetition, so that the post and delayed tests would take approximately 30 minutes each to complete ([Appendix D](#)).

The purpose of using two different question sets on the assessments was to account for the possibility of a testing effect that could confound the results. The testing effect refers to a phenomenon where an initial memory test enhances performance on a later memory test (Chan & McDermott, 2007). In this study, students' completion of the pre-test may have affected what students attended to during the study session and how students performed on the identical questions on the post- and delayed-tests. Thus, a new set of questions (QSB) was presented to

students on the post-test to examine whether or not they would perform differently on new questions compared to questions they had seen before.

*b. Question types*

The questions used on the formal assessments were distributed across four categories originally designed by Chi et al., (1994): 1) Definitions, 2) Explicit, 3) Implicit, and 4) Mental Model questions. Of the total number of questions that appeared on the post and delayed tests, about half from each category were presented on the pre-test (Table 2).

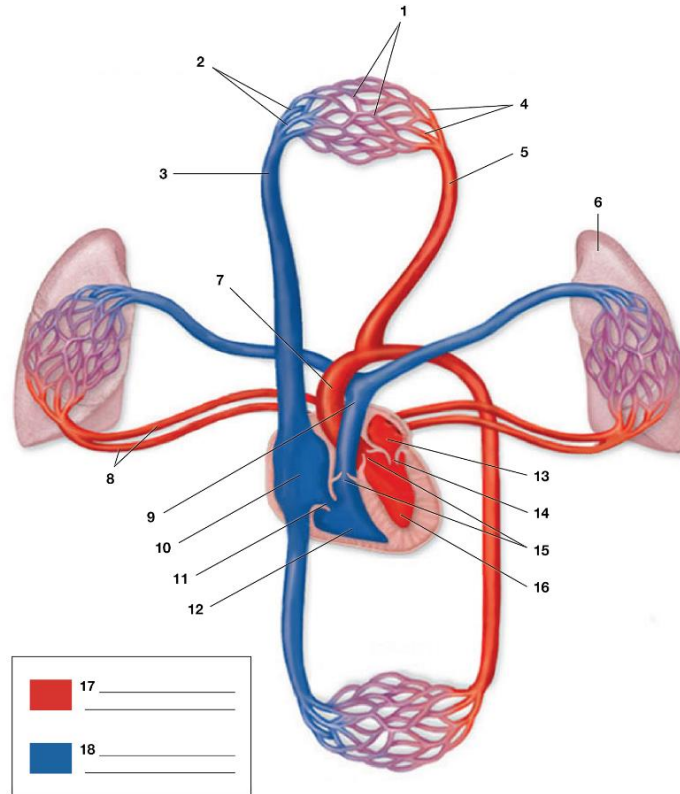
**Table 2:** Types and number of questions on each formal assessment based on Chi et al. (1994)

<i>Assessment</i>	<i>Question Set(s)</i>	<i>Question Type</i>				
		Drawing	Definitions	Explicit	Implicit	Mental Model
<b><i>Pre</i></b>	A	1	5	5	4	4
<b><i>Post</i></b>	A & B	1	10	10	7	7
<b><i>Delayed</i></b>	A & B	1	10	10	7	7

Definition questions required participants to define various components of the CVS including items such as atrium, pulmonary circulation, and septum. Questions in the Explicit category could be answered from information provided explicitly in a line of the CVS text. In order to answer questions in the Implicit category, participants were required to make inferences from the information presented in the text, sometimes across different passages. Finally, Mental Model questions were situated in the larger context of the CVS and aimed at assessing student understanding of the system as a whole. These questions included scenarios where one component in the system was described as altered or damaged, and the student was asked to predict what would be affected by these changes.

*c. Figure interpretation task*

The figure interpretation task was presented to participants at the end of the delayed test. Since the task involved presenting participants with an anatomical diagram of the CVS (Figure 3), it was not presented in earlier assessments in order to not influence participants' drawings. In addition to the diagram presented on the computer, participants were also provided with a colored, paper print-out of the diagram in a plastic sheet protector. They were asked to identify 18 parts of the diagram where indicated and type their answers into a Google Forms page. Then, they were given a dry erase marker and asked to draw the flow of blood wherever they thought it was appropriate.



**Figure 3:** Figure presented to participants for the figure interpretation task.

### 3.6 Data Sources and Reorganization

---

This section describes what data from this study were used for analyses and how they were organized.

#### 3.6.1 Formal assessments

---

Each formal assessment was comprised of a drawing question and free-response questions from the four question categories (Definitions, Explicit, Implicit, and Mental Model). On the final delayed-test, participants also completed a figure interpretation task.

##### *a. Assessment drawings*

The drawing question on each formal assessment appeared first, and asked students to draw a diagram of the CVS using any or all of the three colored pens provided for them. By the end of the study, each participant was expected to have created a drawing of the CVS on three separate occasions: before their study session (on the pre-test), immediately after their study session (on the post-test), and two weeks after that (on the delayed-test).

##### *b. Free-responses*

Participant responses to open-ended questions were collected electronically via Google Forms. After initial collection of the data, two problems were discovered that required some exclusion and reorganization of assessment questions.

First, a few questions were interpreted differently by many students and were removed from the data analysis due to unreliability. For example, based on responses to the question



“Does the blood change in any way as it passes through the heart? If so, how?” some students interpreted “passes through the heart” to mean through one side of the heart while others interpreted it to mean the loop through one side of the heart (to the lungs and back). Thus, when a participant responded that blood gets oxygenated as it passes through the heart, it was difficult to determine whether he or she believed the heart, itself, oxygenates blood or blood gets oxygenated as it loops to the lungs and back to the heart. Consequently, these questions and participant responses to them were not included in the following analyses.

Second, sometimes multiple questions asked about the same content. For example, responses to the Definition question for Valve were often similar to responses to the question that “Why are there valves in the heart?” Given that some participants included information about valves in their response to one of these questions but not the other made it difficult to assess their understanding of valves if responses to these questions were analyzed in isolation. As a result, responses to questions addressing the same content were combined for analyses.

Further restructuring of the free-response data was needed as a result of conflict with the pre-assigned question categories developed by Chi et al., (1994). This restructuring is described in detail in a following section about the analysis of the formal assessments.

*c. Figure interpretation task*

After completing the delayed-test free-response questions, participants were asked to identify various components of an anatomical diagram of the CVS and draw the path of blood flow. Responses were collected in text boxes in Google Forms and the paths drawn on the printed figures were stored as digital photographs.

### 3.6.2 Independent study sessions

---

Various data were collected from independent study sessions. These data will be most informative of how students spent their time while they were left to independently study the CVS using their assigned study strategies.

*a. Time on task*

Day 1 sessions with participants were video and audio recorded. From these videos, different times were measured: total time spent studying, time spent using each study strategy, and the rest of the time (Table 3). Since the experimenter often interacts with the participant as they study Passage 1, this time was excluded from the total independent study time. Instead, the total time spent studying was measured from the time the student began Passage 2 until the final page of the booklet was turned. The time spent using the study strategy was measured as time spent perceptibly writing, speaking aloud, or drawing. The remaining time categorized as “other,” was any time during the study session when the participant was not perceptibly using their study strategy. This time included, but is not limited to, silent thinking, page turning, and re-reading.

**Table 3:** Various times measured from video data

<u>Time spent</u>	<u>Measurement</u>
Studying (total)	Beginning of Passage 2 to the end
Using study strategies	Rewriting, self-explaining, or drawing
Other	Time spent with material, not perceptibly using a study strategy

*b. Self-explanations*

During the independent study sessions, audio and video recorders captured students' self-explanations as they spoke aloud to themselves while reading through the CVS booklet. Each participant in the SE and Combined groups generated self-explanations for each of the 12 passages of text about the CVS. These recordings were transcribed as soon as possible after each Day 1 session.

*c. Booklet drawings*

During the independent study sessions, video cameras recorded participants as they generated drawings. These video data were collected in addition to the drawings themselves. Each participant in the Draw and Combined groups generated a drawing for each of the 12 passages of text about the CVS.

## 4 PART 1: FORMAL ASSESSMENTS

---

Much of the prior research conducted on self-explaining and drawing to learn from text has been done separately (e.g. Chi et al., 1994; Van Meter, 2001) or with substantial scaffolding (e.g. Cromley et al., 2011). This investigation allows for a direct comparison of self-explaining and drawing as independent study strategies for learning about a biological system from text. The goal in analyzing participant responses on formal assessments was to measure their understanding of the CVS before, immediately after, and two weeks after their individual study sessions with the CVS text. This analysis would provide a rough estimate of what learners knew about the CVS prior to entering the study, what they learned from studying the material on Day 1, and what information they retained two weeks after studying.

These assessments were analyzed by their separate parts: the drawing task, the free-response questions, and the figure interpretation task. This section provides descriptions of the analyses conducted on data from each of these parts of the assessments, followed by the results of those analyses.

### 4.1 Drawing Task

---

The examination of student drawings on formal assessments is intended to reveal details of how learners organize information about the CVS in a drawing. Presumed to be some version or extension of a learner's mental model, differences in the quality of drawings (according to the coding scheme) may suggest differences in learning between participants in the various experimental groups.

The drawing task was the first task that participants were asked to complete on each assessment. As described earlier, each participant was provided with colored pens and a blank piece of paper, and asked to draw the CVS. Throughout the study, participants were never shown any figures of the CVS until the very last task on the final day. Thus, their drawings were only influenced by their prior knowledge and what they learned during their study sessions.

#### 4.1.1 Analyses of Drawing Mental Models from Formal Assessment Drawings

---

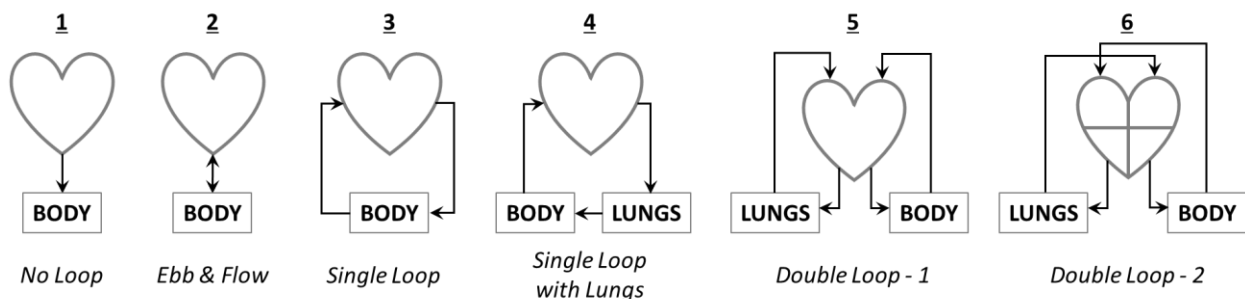
The mental model coding scheme developed in Chi et al. (1994) was used for the initial analysis of the formal assessment drawing data. The coding scheme focused on whether or not participants included complete loops in their systems, and whether or not they correctly included two separate loops for the pulmonary and systemic circulatory sub-systems. In their study, Chi and colleagues assessed student mental models based on answers to a set of their free-response questions. In this current study, however, the mental model coding scheme was used to analyze student drawings.

Drawing, rather than free-response, data were used for mental model analyses for a couple reasons. Coding mental models from free-response data from this study would have roughly equated to participants' total free-response scores because learners with higher scores displayed more knowledge of the subsystem loops. Thus, it would have been somewhat redundant to analyze the free-response data in this manner. Another reason was related to the information presented in the pre-test questions. Many of the questions on the assessments (especially on the pre-test) are quite informative about the CVS. Surely, questions were originally phrased to not be leading, but many of the terms and some of the wording of questions unavoidably presented new information to students, allowing them to learn about the CVS simply by reading the pre-test questions. For example, by reading the question "Where does

blood go to after it's oxygenated?" a participant can infer that blood gets oxygenated at some point in the system, blood travels to a particular place in the body to get oxygenated, blood travels to a particular place after it is oxygenated, and the system contains oxygenated and deoxygenated blood separately. Informative questions such as this one, made it difficult to distinguish in free-response answers what information was genuinely part of the student's mental model and what was simply being restated as part of the question. The drawing task, which was presented as the very first task on each assessment (before participants saw any free-response questions), did not have this problem.

Before describing the coding scheme for identifying models from these drawings, it must be noted that what can be extracted from learners' drawings are not complete pictures of their mental models of the system. The drawings will indeed be reflective of the student's conceptualization of the CVS, but they will not contain all of the information that exists in the learner's mental model of the system. In particular, many behaviors and functions will not be represented in these drawings but are key parts of their internal mental models. Therefore, this glimpse or type of mental model being identified will specifically be referred to as a "drawing mental model (DMM)" to distinguish it from a true mental model.

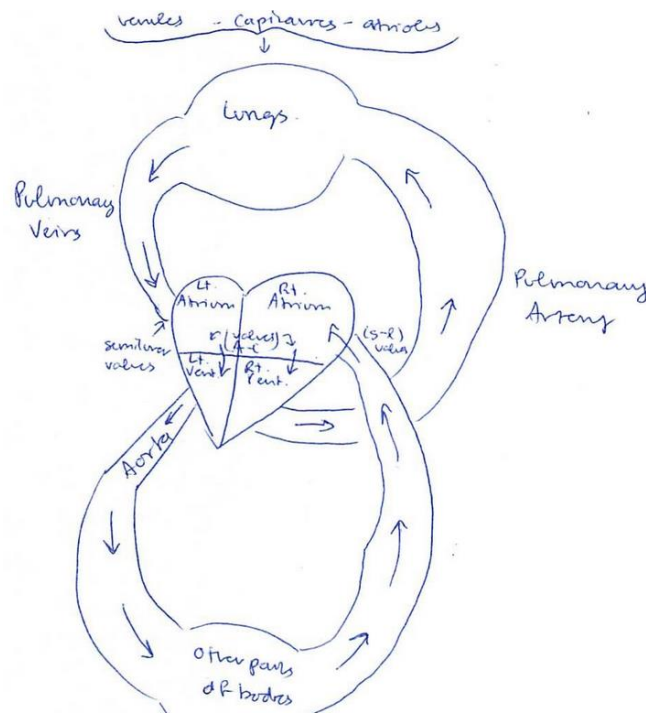
The six mental models identified in Chi et al. (1994) were used to code each student drawing on a 6-point scale. Representations of these models are shown in Figure 4. In a No Loop model (1 point), blood is drawn leaving the heart but does not return. In an Ebb and Flow model (2 points), blood exits and returns to the heart but via the same blood vessel. In a Single Loop model (3 points), blood exits and returns to the heart via different blood vessels. In a Single Loop with Lungs model (4 points), blood exits and returns to the heart including the lungs somewhere along the loop. In a Double Loop-1 model (5 points), blood flows to the lungs and back to the heart, and also to the body and back to the heart, but the drawing lacks details necessary to qualify as a Double Loop-2 model. In a Double Loop-2 model (6 points), blood flows in two separate loops to the lungs and the body and the drawing also contains correct specifics about how blood flows through the four chambers of the heart.



**Figure 4:** The six mental models of the cardiovascular system identified by Chi et al (1994). Figure adapted from Friedman & Forbus (2011).

The only deviation in this analysis from the mental model coding scheme developed by Chi *et al* (1994) was excluding one coding element that required students to describe how the lungs play a role in the oxygenation of blood in the Single Loop with Lungs (4 points) and Double Loop-1 (5 points) models. Had this coding scheme been used to analyze free-response data, the assessment questions may have prompted participants to address this item, specifically. With drawings, however, participants often depicted lungs and blood vessels passing through them but did not always explicitly show that the lungs oxygenated this blood (e.g., Figure 5). Participants who drew the CVS in this way likely assumed that the association between the lungs

and oxygen was sufficient for a viewer to interpret that blood in these vessels becomes oxygenated in the lungs.



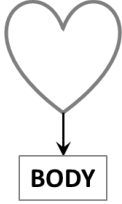
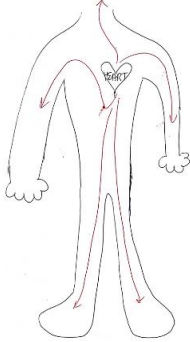
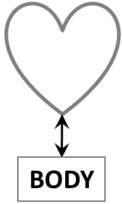
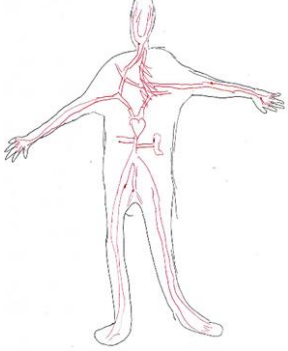
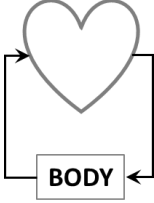
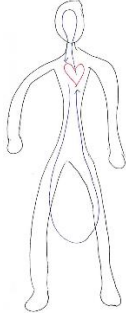
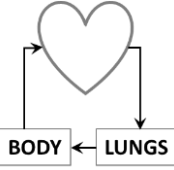
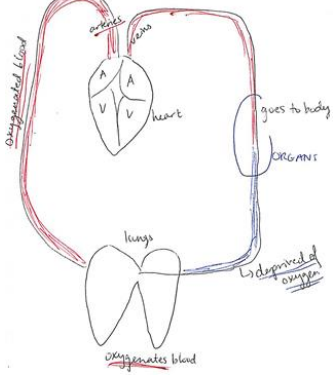
**Figure 5:** Example drawing that does not indicate that the lungs play a role in the oxygenation of blood (#0491 Post-test)

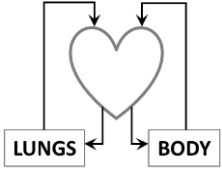
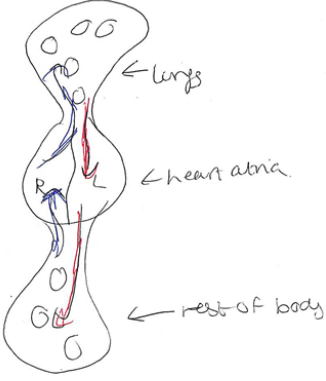
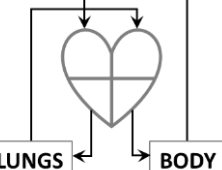
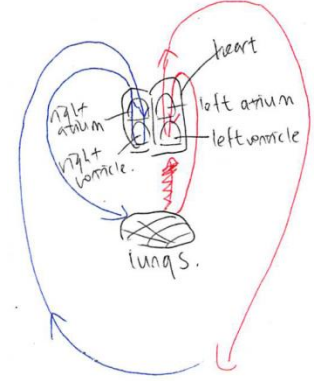
With the original coding scheme, indicating that the lungs play a role in the oxygenation of blood was necessary for a drawing to qualify as any of the three highest levels of mental models. In other words, if a drawing clearly included double loops but lacked this item, it could only be coded as is a Single Loop model (3 points), even if lungs were included in the drawing and the remaining details were correct, such as in Figure 5.

By excluding this item in the modified coding scheme, students who created drawings with double loops (one passing through the lungs and one passing through the body) would be assumed to know that the lungs play a role in the oxygenation of blood. Thus, these types of drawings were coded as Double Loop-1 models (5 points) because they included all of the necessary features of a double loop but lacked this detail (and perhaps others) to qualify as a Double Loop-2 model (6 points). Student examples for each mental model are provided in Table 4, including the list of necessary features that qualified each drawing for that model type.

**Table 4:** Mental model coding scheme for drawings on formal assessments.

Mental Model	Example Drawing	Necessary Features for Each Type of Mental Model from Chi <i>et al</i> (1994)
--------------	-----------------	---

1	<p>No Loop</p> 	 <p>(#5976 Pre-test)</p>	<ul style="list-style-type: none"> <li>• Blood is pumped from the heart to the body.</li> <li>• Blood does not return to the heart.</li> </ul>
2	<p>Ebb and Flow</p> 	 <p>(#3932 Pre-test)</p>	<ul style="list-style-type: none"> <li>• Blood is primarily contained in blood vessels.</li> <li>• Blood is pumped from the heart to the body.</li> <li>• Blood returns to the heart by way of the same blood vessel.</li> </ul>
3	<p>Single Loop</p> 	 <p>(#5737 Pre-test)</p>	<ul style="list-style-type: none"> <li>• Blood is primarily contained in blood vessels.</li> <li>• Blood is pumped from the heart to the body.</li> <li>• Blood returns to the heart from the body.</li> </ul>
4	<p>Single Loop with Lungs</p> 	 <p>(#0502 Pre-test )</p>	<ul style="list-style-type: none"> <li>• Blood is primarily contained in blood vessels.</li> <li>• Heart pumps blood to body or lungs.</li> <li>• Blood returns to heart from body or from lungs.</li> <li>• Blood flows from lungs to body or from body to lungs without return to heart in between.</li> </ul>

5	<p>Double Loop-1</p> 	 <p>(#0432 Pre-test)</p>	<ul style="list-style-type: none"> <li>• Blood is primarily contained in blood vessels.</li> <li>• Heart pumps blood to body.</li> <li>• Blood returns to heart from body.</li> <li>• Heart pumps blood to lungs.</li> <li>• Blood returns to heart from lungs.</li> </ul>
6	<p>Double Loop-2</p> 	 <p>(#8712 Pre-test)</p>	<ul style="list-style-type: none"> <li>• All features from Double Loop-1.</li> <li>• Heart has four chambers</li> <li>• Septum divides heart lengthwise</li> <li>• Blood flows from top to bottom.</li> <li>• At least 3 of the following: Blood flows from... <ul style="list-style-type: none"> <li>○ Right ventricle to lungs</li> <li>○ Lungs to left atrium</li> <li>○ Left ventricle to body</li> <li>○ Body to right atrium</li> </ul> </li> </ul>

Note that this 6-point scale does not reflect what is expected to be a non-linear progression for learning about the CVS. That is to say, an improvement from a Single Loop model (3 points) to a Single Loop with Lungs model (4 points) may be easier to accomplish than an improvement from a Single Loop with Lungs model (4 points) to a Double Loop-1 model (5 points), in terms of learning. However, this scale allows us to simply transform the drawings into quantitative data to compare performance between groups.

#### 4.1.2 Formal Assessment Drawing Results

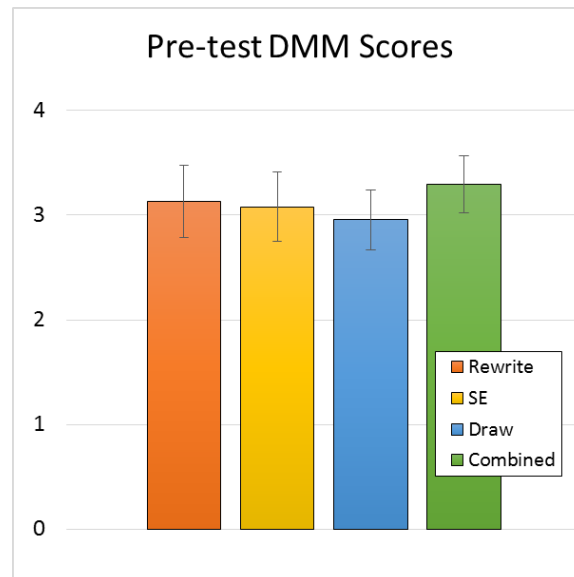
Two raters independently coded each participants' drawings on the three assessments for mental model type. Cohen's  $\kappa$  was calculated to determine the amount of agreement between the two raters on scoring drawings on each assessment. There was very good agreement between the two raters on each assessment: pre-test  $\kappa = 0.85$ , post-test  $\kappa = 0.83$ , and delayed test  $\kappa = 0.81$ .

One pre-test drawing and two delayed-test drawings were missing from the data set. Two post-test drawings were excluded because they are believed to have been drawn hastily, based on the quality of their drawings and a comparison of the times these specific participants spent on their drawings. Average unadjusted drawing scores for each assessment are presented in Table 5, below.

**Table 5:** Data summary of drawing mental model scores (and standard deviation) by assessment, based on a 6-point scale.

	Pre-test		Post-test		Delayed-test	
	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)
Rewrite	24	3.13 (1.68)	24	5.04 (1.23)	24	4.92 (1.02)
SE	25	3.08 (1.66)	25	4.12 (1.72)	24	3.96 (1.68)
Draw	22	2.95 (1.33)	22	5.32 (0.89)	22	4.86 (1.17)
Combined	24	3.29 (1.33)	23	5.26 (0.54)	24	5.00 (0.66)

A one-way ANOVA was used to determine if mean pre-test DMM scores were statistically different across the four groups. Scores were not normally distributed by group, but a one-way ANOVA is robust enough to tolerate a violation of normality well. There was homogeneity of variances, as assessed by Levene’s test of homogeneity of variances ( $p = 0.134$ ). The ANOVA determined that the differences between the average DMM scores across groups on the pre-test were not statistically significant (Figure 6).

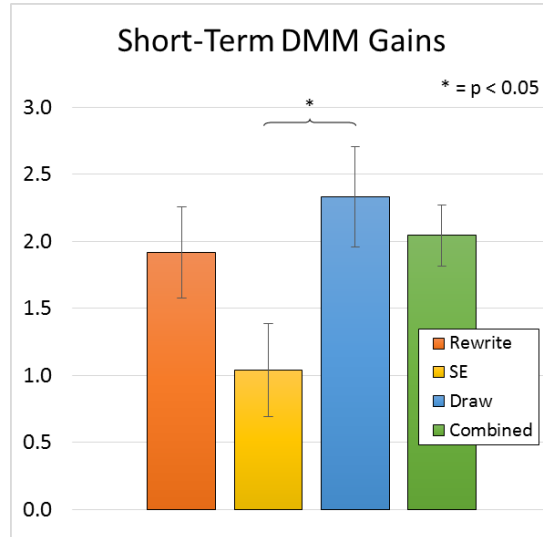


**Figure 6:** Mean drawing mental model pre-test scores with standard errors.

ANCOVA diagnostics of pre and post-test DMM scores revealed that the data violated some of the assumptions required to accurately conduct an ANCOVA. Therefore, the data was transformed by subtracting pre-test DMM scores from post-test DMM scores to give us *short-term DMM gain scores*, and a one-way ANOVA was conducted with these gain scores. These scores were normally distributed and ANOVA tests are robust enough to tolerate violations of normality. There was homogeneity of variances, as assessed by Levene’s test of homogeneity of variances ( $p = 0.126$ ). Paired t-tests determined that increases from pre to post-test DMM scores for each condition were significant (Rewrite group:  $t(23) = 5.64, p < 0.001$ ; SE group:  $t(24) = 3.02, p < 0.01$ ; Draw group:  $t(20) = 6.24, p < 0.001$ ; Combined group:  $t(22) = 8.85, p < 0.001$ ).



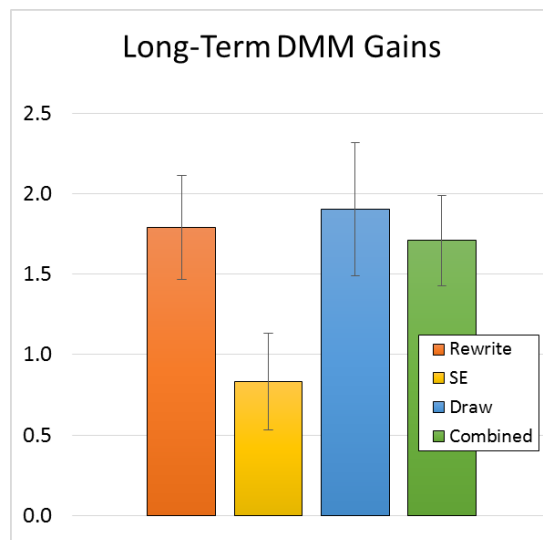
The one-way ANOVA of short-term DMM gain scores determined that study strategy had a significant effect on gains,  $F(3,89) = 2.812$ ,  $p \leq 0.05$ , partial  $\eta^2 = 0.087$ . The SE group had the least amount of gains of all four conditions. Post hoc t-tests with Bonferroni adjustments determined that the SE group gained significantly less in DMM scores than the Draw group ( $p \leq 0.05$ ; Figure 7).



**Figure 7:** Mean short-term drawing mental model gain scores with standard errors.

For the same reasons that post-test DMM scores were transformed, an analogous transformation was conducted to generate *long-term DMM gains* by subtracting pre-test DMM scores from delayed-test DMM scores. The distribution of these gain scores were skewed to the left but ANOVA tests are robust enough to tolerate violations of normality. There was homogeneity of variances, as assessed by Levene’s test of homogeneity of variances ( $p = 0.366$ ).

The one-way ANOVA test determined that differences in long-term DMM gains across groups were not significantly significant (Figure 8). However, the pattern across groups was nearly identical to that seen with the short-term DMM gains scores (Figure 7).



**Figure 8:** Mean long-term mental model drawing gain scores with standard errors.

Taken together, these results suggest that the SE group learned the least about drawing the CVS, compared to the other three groups.

## 4.2 Free-response Questions

---

After completing the drawing task portion of the formal assessment, participants were presented with free-response questions. This section explains problems with using coding schemes developed in previous studies to analyze the responses in this study, and details the development of a different scheme. Following that is an explanation of adjustments made to account for differences between groups in pre-test scores at the beginning of the study and the results of group comparisons after controlling for these scores.

### 4.2.1 Analysis of Free-Response Questions

---

The categorization of assessment questions developed by Chi et al. (1994), was found to be problematic for analyses of free-response data in this study. Many questions did not fit cleanly into individual categories and multiple questions often overlapped in topic, asking about the same CVS component in slightly different ways. Furthermore, the method of analysis used by Ainsworth & Loizou (2003) was also problematic because it focused more on the CVS text and less on inferences that students could make to improve their understanding. The solution was to develop a new coding scheme based on a structure-behavior-function framework and to account for questions that asked about the path of blood flow with an additional category. The following details the problems found with the previous studies' coding schemes, the development of a new one, and how assessment questions were re-categorized to fit this new scheme.

#### *a. A consideration of coding schemes used in previous studies*

Two methods of analyses from prior studies were considered. Given that the questions on the formal assessments were taken from Chi et al. (1994), the method of analysis used in that study was considered first. Chi and colleagues developed a coding scheme that considered responses to each assessment question separately and by question category (i.e., Definition, Explicit, Implicit, and Mental Model). Each response was assigned up to a maximum of six points each. While the questions used in this study were identical to those in Chi et al. (1994), applying this method of analysis to the data posed three problems. First, some assessment questions targeted different knowledge categories while others did not. For example, Definition questions often elicited responses about structures, behaviors, and functions, while Explicit questions only sometimes elicited responses about structures. As a result, if six points were attributed to each question, those individual points would seemingly be worth more or less (in terms of representing student understanding) depending on which and how many knowledge categories the question addressed.

Second, and closely related to the first point, while the assessment questions were specially designed as four categories of questions that increased in difficulty, six points were assigned to each question in all categories. In other words, all questions were weighed equally, regardless of the difficulty of the question category it was in. As a result, total scores would completely ignore the fact that the questions fit into categories of ascending difficulty. Furthermore, even totals within categories would ignore the differences in difficulty or differences in learner understanding that is elicited between individual questions.

Third, multiple questions sometimes assessed student understanding of the same content. Consequently, analyzing student responses to each question in isolation would not account for

instances when a student clearly demonstrated their understanding of a concept in response to one question but chose not to repeat the information in their response to another that assessed the same or similar idea. For these reasons, the coding scheme used in Chi et al. (1994) was not used in this study.

The coding scheme developed and used in Ainsworth & Loizou (2003) was also considered for use with the formal assessment data, but was also found to be unsuitable for our purposes. In that study, the researchers administered a subset of the assessment questions used in Chi et al. (1994) but only used what was explicitly stated in the CVS text as the sole source for possible correct answers. Essentially, they created a checklist for each assessment question by listing specific lines in the CVS text that could be used to appropriately answer each question. Responses were then given differing amounts of points for each question based on the difficulty of the question and how many items from the checklist were included in the response. A benefit of using this method of analysis was that it systematically considered all of the information that participants could have learned from the CVS text. It also weighed questions differently across the question categories based on how much information could be elicited in a response to that question. However, the strict references to individual lines in the CVS text place an emphasis on specifics that students could remember from the text, rather than on the integration of the information and the generation of inferences.

Given that the Implicit and Mental Model questions were designed to specifically require students to make inferences about the text, it would be difficult to map the correct answers to these questions onto specific lines in the CVS text. For example, one assessment question asks students to explain why the right side of the heart is less muscular than the left. While the text describes where blood flows to and from on each side of the heart, the concepts of pressure, force, and distance are never used. Yet many students correctly explain that the distance from the right side of the heart to the lungs is much shorter than the distance from the left side to the furthest cell in the body. Some students even explain that, as a result of this difference in distance, a greater pumping force is needed or a greater amount of pressure must be generated in order to get blood to reach all of the cells in the body. These ideas cannot be mapped onto specific lines in the CVS text. Therefore, this method of analysis was not used in this study.

For purposes of this research project, a new coding scheme was developed, greatly influenced by Chi et al. (1994) and Ainsworth & Loizou (2003), and attempted to account for the problems described above. Rather than assessing students based on their specific references to lines in the CVS text, the new coding scheme aimed to account for all pieces of knowledge related to pathways, structures, behaviors, and functions that students displayed in their responses to questions.

*b. Structure-Pathway-Function coding scheme*

A priority for developing a new coding scheme for this data was to use patterns found in student responses to free-response questions. The a priori design of the four categories of questions by Chi et al. (1994) was problematic because *responses* to these questions did not reflect these distinct categories. Based on free-response data, these four categories did not necessarily ascend in difficulty as they were intended to in the original study, and some assessment questions could fit into more than one of these category. For example, the Explicit question asking about valves elicited responses similar to those for the Definition question for valves. Instead, the Structure-Behavior-Function (SBF) model was considered as a framework for analyzing student knowledge.

The SBF model is a useful framework for studying how learners organize their knowledge of complex systems. It can account for differences between novices and experts in their understanding of complex systems in biology, and has been suggested to allow learners to see how different components and features contribute to the system (see Hmelo-Silver & Pfeffer, 2004). The structure category refers to information about the physical characteristics of a component within the system, behavior describes the activity of the component, and function refers to the purpose of that component in the system. For example, the structure of an atrium is a muscular chamber; its behavior is that it contracts and relaxes, squeezing blood; and one of its functions is to serve as a holding chamber for blood before it enters the ventricle (Chi et al., 1994). These SBF categories can help learners to organize their knowledge of the components in the CVS and the relationships between those components.

While the SBF framework is particularly useful when investigating learner understanding of biology content, it was challenging when applied to this study, especially based on the assessment questions that we used. This is because, while structures are relatively simple to identify in the system, behaviors and functions are often difficult to identify and distinguish from each other. For example, an artery is a blood vessel that carries blood away from the heart. Its structure is a hollow tube of three layers of tissue and its function is to carry blood in a direction away from the heart. However, given that an artery is a passive vessel that holds the blood that gets pumped through it, it is difficult to identify its behavior in this context.<sup>1</sup> Similar issues surfaced about what qualified as behaviors versus functions when considering other types of blood vessels and subsystems (i.e., systemic and pulmonary circulation) that were items in the definition section of the assessments.

When examining assessment data, student responses to definition questions mostly described structures and functions of the components without describing behaviors. Higher-level assessment questions (e.g., Implicit and Mental Model questions) tended to begin with “why,” rather than “how,” often eliciting responses that would describe functions but not behaviors. These questions did not provide participants with enough prompting to expect them to describe all three SBF categories (or at least did not do so consistently). Thus, the SBF model needed to be amended to better fit the data in this study.

The SBF coding scheme was amended to account for the difficulty in identifying behaviors for components within the CVS and in distinguishing behaviors from functions. A category for *structures* was still used, and still referred to physical characteristics of parts within the system. The category for behaviors was absorbed by the category for functions, so that all responses that described actions or purposes of components within the CVS were simply considered descriptions of *function*. Given that the function of some components are in fact to behave in a certain manner, we kept the word “function” to label this knowledge category. Finally, a new category for the understanding of *pathways* was added to this scheme.

Many assessment questions asked learners where blood travels to or from at various points in the CVS, and responses to these types of questions rarely described structures, behaviors, or functions. The pathways category accounted for responses to these types of questions. For this particular subject domain, understanding the sequence of organs and tissues that blood travels through is a crucial part of understanding the CVS, and may be a type of knowledge that is different from structures and functions. Structures focus on the physical form

---

<sup>1</sup> Arterial walls do have a behavior that is elastic in nature to accommodate changes in blood pressure, but this elasticity is not mentioned or implied in the CVS text, and is not necessary for understanding the role of arteries in this context.

of components, pathways focus on the sequence of components that blood travels through, and functions focus on the activity and purpose of components. Knowledge of each category is useful for connecting to other knowledge within and across categories in order to understand how the components work together in the system.

In terms of a structures-*pathways*-functions (SPF) model, assessment questions that allowed students to demonstrate an understanding of structures often asked them to provide a description of a CVS component based on its name, or to name a component based on its description. Understanding of pathways was examined in responses about the order in which that blood travels to and from organs and tissues in the system. Such responses generally described or named components that came before or after another in the sequence. Lastly, functions referred to the actions and roles of parts of the system, requiring an understanding of how a component acted or why it was a necessary part of the system. This type of understanding was reflected in responses to questions that asked about why an action was most efficient or what would happen to the system if one part was dysfunctional.

*c. Re-categorization of assessment questions*

Since assessment questions administered in this study were organized by the four question categories developed by Chi et al. (2001), individual questions needed to be reevaluated to determine whether it allowed learners to display knowledge of structures, pathways, and/or functions. Most items in the Definitions category tended to elicit responses that described both the structure and function of the item. For example, when participants were asked to provide a definition for Artery, many described that it is a blood vessel (structure) that carries blood away from the heart (function).

Three questions in the Explicit category were combined with Definition questions because they referred to the same content as a Definition question. For example, most responses to the definition of Valve paralleled responses to the Explicit question, “Why are there valves in the heart?”<sup>2</sup> In such instances, when two questions asked about the same CVS component, responses to both questions were combined and analyzed as one response. The analysis evaluated whether or not the combined responses correctly described the structure and function of the item. The remaining questions in the Explicit category mostly focused on student understanding of the pathway that blood travels as it is pumped throughout the body (e.g., “Where does blood go after it’s oxygenated?”).

A majority of questions from the Implicit and Mental Model categories addressed student understanding of functions. For example, from the Implicit category, a question asked about the function of capillaries: “Why are capillary walls one-cell thick?” and from the Mental Model category a question asked about the function of the septum: “What is the consequence at the cellular level of having a hole in the septum?”

Overall, QSA (on the pre, post, and delayed tests) contained 18 questions, and QSB (on the post and delayed tests) contained 16 questions. Thus, the post and delayed tests each contained 34 questions total. Responses to two questions from QSA and two questions from QSB were excluded from analyses due to the unreliability explained above (Section 3.6.1b). Responses to the remaining 30 questions were analyzed to determine when each provided an opportunity for a student to display knowledge of structures, pathways, and/or functions.

---

<sup>2</sup> In this example, the Explicit question asks specifically about valves in the heart, while the Definition question asks about valves in general. Given that learners very rarely made this distinction in their responses to both questions, it was reasonable to combine them for analyses.

QSA provided students with 7 opportunities to display their understanding of structures, 4 opportunities for pathways, and 13 opportunities for functions. QSB provided students with 8 opportunities for students to display knowledge of structures, 4 opportunities for pathways, and 11 opportunities for functions (Table 6). Hence, the post and delayed tests provided students with a total of 15 opportunities to display their understanding of structures, 8 opportunities for pathways, and 24 opportunities for functions. Details of opportunities to display SPF knowledge are presented in [Appendix D](#).

**Table 6:** Number of opportunities for participants to display SPF knowledge on QSA and QSB free-response questions

<b>Number of Opportunities to Display Knowledge</b>			
<i>Knowledge Category</i>	Question Set A	Question Set B	<b>Total</b>
<b><i>Structures</i></b>	7	8	15
<b><i>Pathways</i></b>	4	4	8
<b><i>Functions</i></b>	13	11	24
<b>Total</b>	24	23	47

Note that some individual questions allowed participants to mention more than one structure, pathway, or function. For example, the question “What keeps blood flowing in a given direction when it leaves the heart?” has two correct answers: valves and pressure differentials. These types of questions were counted as two opportunities to display knowledge about functions. It can be inferred that a student who provided a complete response to this question understood that the functions of valves and pressure differentials are to keep blood flowing in one direction. Accordingly, each student’s response to each question was analyzed for whether or not they correctly identified or described structures, pathways, and/or functions where these opportunities were identified.

#### 4.2.2 Free-Response Results

Two raters independently coded each participants’ free-response answers for correct mentions of structures, pathways, or functions where identified in the coding scheme. Cohen’s  $\kappa$  was calculated to determine the amount of agreement between the two raters on each assessment. There was very good agreement between the two raters on each assessment, pre-test  $\kappa = 0.94$ , post-test  $\kappa = 0.95$ , and delayed-test  $\kappa = 0.96$ .

Overall raw scores were calculated by totaling the number of times a student correctly identified a structure, pathway, or function. Percentages were then calculated based on the total number of opportunities students had to display SPF knowledge on each Question Set (see Table 6 for total number of opportunities). Average unadjusted percentage scores for each Question Set on each assessment are shown in Table 7.

In regards to most analyses of statistical differences between groups, scores on the various components of the formal assessment were compared using analyses of covariance (ANCOVA) tests in order to determine whether or not significant differences existed between

groups after using the different study strategies.<sup>3</sup> Furthermore, using pre-test scores as a covariate helped us to account for some differences that existed between students when they entered the study.

The following are results from these analyses of the free-response data. Results from analyses of QSA are presented first, in the order of the three assessments (pre, post, then delayed), followed by the results from QSB. The pre-test scores are presented as unadjusted percentages, but post and delayed test scores are presented as adjusted percentages as a result of the ANCOVAs conducted. This section concludes with a summary of the results, including a figure that combines the SPF scores from each question set on the post and delayed tests (Figure 14).

---

<sup>3</sup> Given that the Rewrite group ultimately was not a control group and scores were compared across all four conditions, it was more appropriate to conduct ANCOVA tests followed by post hoc t-tests, rather than only using t-tests to reduce the likelihood of committing a type I error. Moreover, rather than using regression analyses to test whether or not the study strategies had an effect on performance, we used ANCOVA analyses to see if the effects of study strategies on performance are statistically different across groups.

**Table 7:** Data summary of unadjusted mean scores (and standard deviations) for each group on each formal assessment. Scores are each presented as a percentage of the total possible points for that overall score or sub-score. Sub-scores: S = Structures, P = Pathways, F = Functions.

Unadjusted Question Set A Scores

Condition	PRE-TEST (%)				POST-TEST (%)				DELAYED-TEST (%)			
	Overall	S	P	F	Overall	S	P	F	Overall	S	P	F
<b>Rewrite</b> (n=24)	35.76 (20.63)	37.50 (21.40)	32.29 (29.93)	35.90 (21.24)	67.53 (15.97)	88.69 (11.90)	61.46 (25.52)	58.01 (21.03)	65.1 (17.5)	83.33 (18.68)	62.50 (26.58)	56.09 (20.35)
<b>SE</b> (n=25)	38.00 (18.95)	36.57 (22.99)	33.00 (22.50)	40.31 (20.98)	63.00 (18.21)	71.43 (22.21)	59.00 (27.84)	59.69 (20.02)	60.17 (21.18)	68.57 (27.66)	50.00 (31.46)	58.77 (19.34)
<b>Draw</b> (n=23)	34.78 (20.16)	32.92 (22.97)	27.17 (31.00)	38.13 (18.77)	71.38 (14.29)	88.82 (12.89)	71.74 (20.37)	61.87 (18.04)	63.95 (16.22)	76.40 (21.80)	61.96 (23.68)	57.86 (20.07)
<b>Combined</b> (n=24)	36.28 (19.17)	44.05 (25.25)	23.96 (27.07)	35.90 (17.81)	69.27 (16.93)	83.93 (18.51)	66.67 (26.24)	62.18 (19.63)	67.36 (19.72)	79.17 (23.82)	60.42 (26.50)	63.14 (21.52)

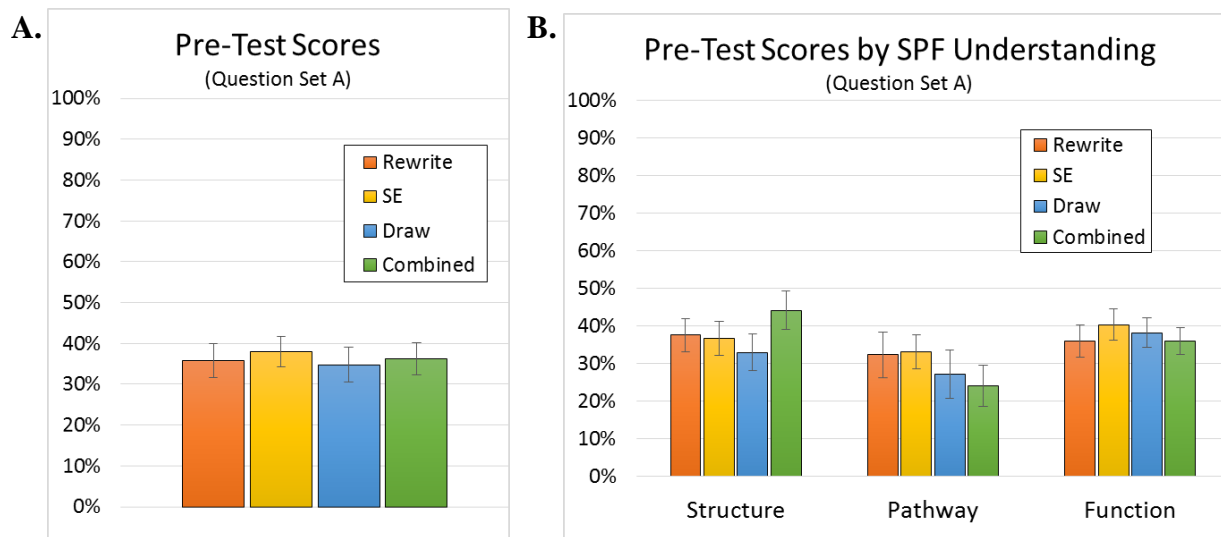
Unadjusted Question Set B Scores

	POST-TEST (%)				DELAYED-TEST (%)			
	Overall	S	P	F	Overall	S	P	F
<b>Rewrite</b> (n=24)	61.96 (16.18)	74.48 (18.24)	68.75 (18.43)	50.38 (18.37)	57.61 (18.37)	67.19 (17.99)	63.54 (23.29)	48.48 (22.54)
<b>SE</b> (n=25)	56.52 (16.51)	64.50 (19.66)	66.00 (21.51)	47.27 (17.41)	58.43 (21.04)	61.00 (26.84)	59.00 (24.87)	56.36 (21.80)
<b>Draw</b> (n=23)	62.57 (16.15)	74.46 (15.30)	71.74 (25.34)	50.59 (22.07)	58.41 (15.99)	63.59 (18.81)	65.22 (20.97)	52.17 (21.62)
<b>Combined</b> (n=24)	63.59 (18.3)	73.44 (19.96)	72.92 (20.74)	53.03 (22.05)	62.5 (16.85)	70.83 (20.74)	68.75 (18.43)	54.17 (18.08)



a. *Question Set A: Pre-assessment scores*

Mean overall pre-test scores for each condition are presented in Figure 9A. A one-way multivariate analysis of variance (MANOVA) was used to determine if there were significant differences in pre-test scores across the four study strategy conditions. The analysis was conducted with pre-test sub-scores (structures, pathways, and functions) as the dependent variables. Examinations of possible violations of MANOVA assumptions are presented in [Appendix E](#). The analysis determined that overall pre-test scores were not statistically different across the four conditions,  $F(9,219) = 1.428, p > 0.05$ ; Wilks'  $\Lambda = 0.871$ ; partial  $\eta^2 = 0.045$ . Differences in pre-test SPF sub-scores were also not statistically significant (Figure 9B).

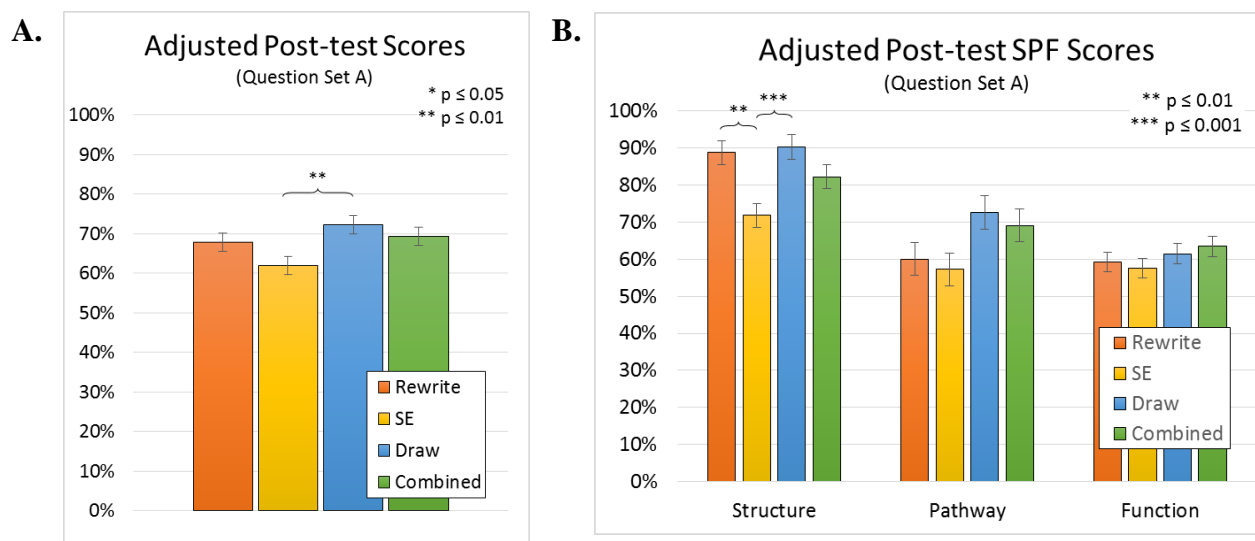


**Figure 9:** **A.** Mean pre-test scores (QSA) with standard errors. **B.** Mean pre-test (QSA) sub-scores by structure, pathway, and function with standard errors.

Although these pre-test scores were not statistically different, one-way analyses of covariance (ANCOVAs) were conducted on subsequent comparative tests to account for the minor differences that did exist. In other words, pre-test scores were controlled for when differences between post- and delayed-test scores were examined.

b. *Question Set A: Post-assessment scores*

By comparing post-test scores, while controlling for pre-test scores, inferences can be made about what students learned from their study sessions. Examinations of possible violations of assumptions for the ANCOVAs conducted in this section and adjusted mean scores are presented in [Appendix F](#). After controlling for overall pre-test scores, study strategy was found to have a significant effect on overall post-test QSA scores,  $F(3, 91) = 3.515, p \leq 0.01$ , partial  $\eta^2 = 0.104$ . Post hoc t-tests, performed with Bonferroni adjustments, revealed that overall post-test QSA scores were significantly lower for students in the SE group than those in the Draw group ( $p \leq 0.01$ ; Figure 10A).



**Figure 10: A.** Adjusted mean post-test QSA scores with standard errors. **B.** Adjusted mean post-test QSA SPF sub-scores with standard errors. Asterisks indicate significant differences between groups, determined by post hoc t-tests.

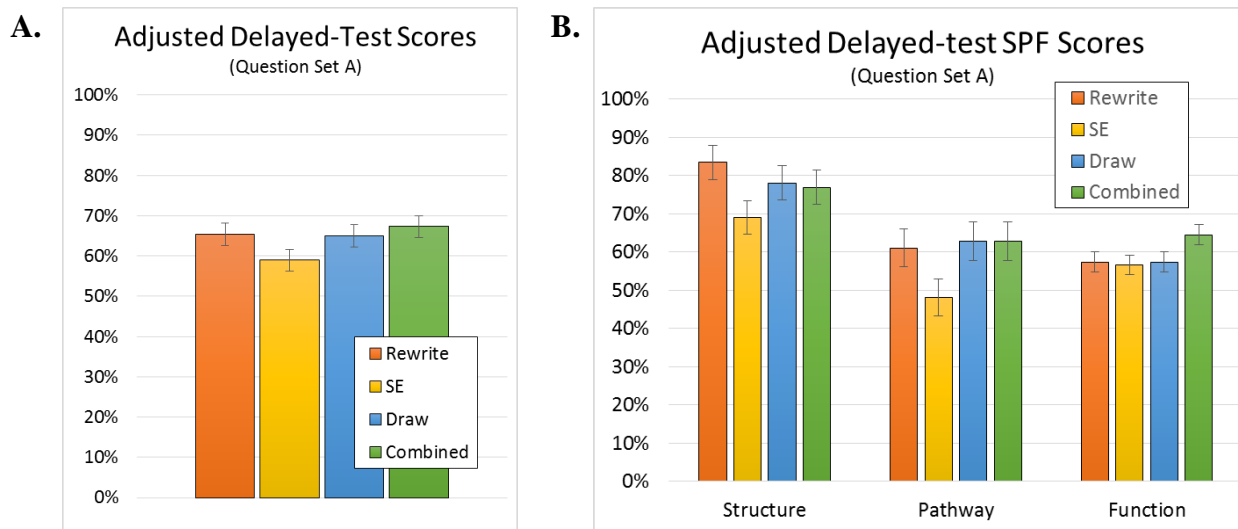
Differences found between overall post-test QSA scores motivated a closer investigation of knowledge categories by groups. After controlling for pre-test structures sub-scores, study strategy was found to have a significant effect on post-test QSA structure sub-scores,  $F(3, 91) = 6.851, p \leq 0.001$ , partial  $\eta^2 = 0.184$ . Post hoc t-tests (with Bonferroni adjustments) determined that average post-test QSA structures sub-scores were significantly lower for students in the SE group than the Rewrite group ( $p \leq 0.01$ ) and the Draw group ( $p \leq 0.001$ ; Figure 10B, left). After controlling for pre-test pathways scores, there was also a statistically significant difference in average post-test QSA pathways scores across groups,  $F(3, 91) = 2.674, p \leq 0.05$ , partial  $\eta^2 = 0.081$ , but post hoc t-tests did not reach statistical significance. However, the SE group had the lowest mean sub-score for this category while the Draw group had the highest (Figure 10B, center). Finally, after controlling for pre-test functions sub-scores, the differences between average post-test QSA functions sub-scores were not found to be statistically significant (Figure 10B, right).

*c. Question Set A: Delayed-assessment scores*

QSA scores on the delayed test were also compared across groups, using pre-test scores as covariates. Examinations of possible violations of assumptions for the ANCOVAs conducted in this section and adjusted mean scores are presented in [Appendix G](#). These ANCOVAs and post hoc tests comparing QSA delayed-test scores between groups were not statistically significant.

After adjusting for overall pre-test scores, the mean overall delayed-test QSA scores for participants in the SE group was, again, the lowest of the four groups but was not statistically different from the other groups (Figure 11A). A closer look at delayed-test scores by SPF knowledge categories is shown in Figure 11B. After controlling for respective pre-test sub-scores, differences in delayed-test SPF sub-scores examined by ANCOVAs were not statistically

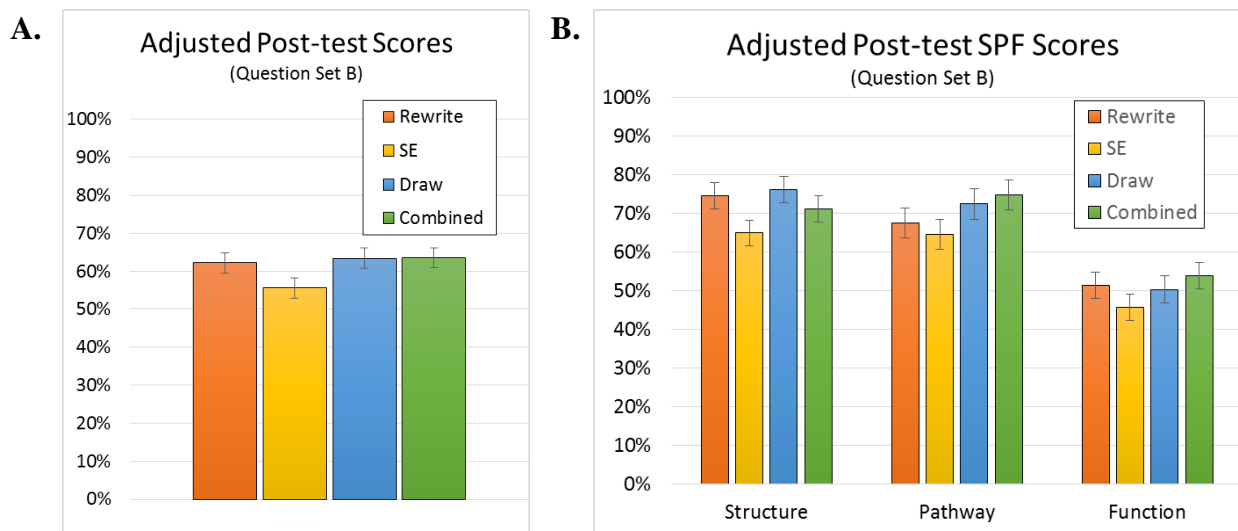
different. However, the mean sub-scores for the SE group were, again, the lowest of the four groups for each of the three SPF sub-scores.



**Figure 11:** **A.** Adjusted mean delayed-test QSA scores with standard errors. **B.** Adjusted mean delayed-test QSA SPF sub-scores with standard errors.

*d. Question Set B: Post-assessment scores*

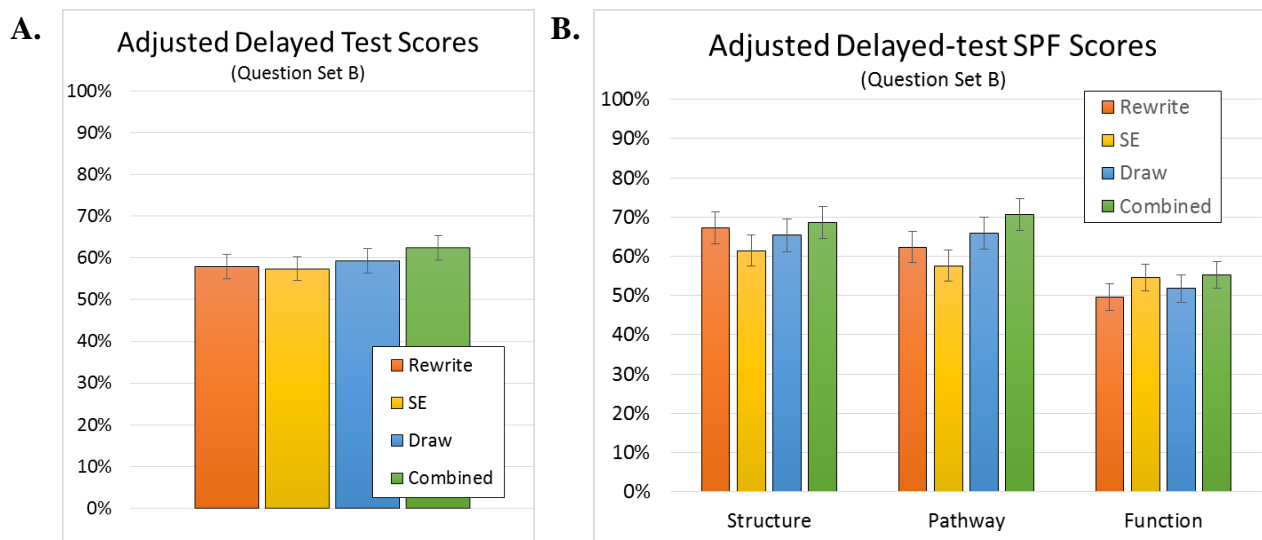
QSB did not appear on the pre-test, but pre-test (QSA) scores were used as covariates for ANCOVAs conducted with QSB scores since they provided rough estimates of participants' initial SPF understanding. Examinations of possible violations of assumptions for the ANCOVAs conducted in this section and adjusted mean scores are presented in [Appendix H](#). ANCOVAs conducted with QSB scores revealed patterns similar to those seen with QSA but none of these tests reached statistical significance. Figure 12A shows that participants in the SE group had the lowest mean score of the four groups, while the other three were very similar. ANCOVAs conducted on post-test SPF sub-scores, after controlling for respective pre-test sub-scores, were not statistically significant, but mean sub-scores for the SE group were, again, the lowest of the four groups for each of the SPF categories (Figure 12B).



**Figure 12:** **A.** Adjusted mean post-test QSA scores with standard errors. **B.** Adjusted mean post-test QSB SPF sub-scores with standard errors.

e. *Question Set B: Delayed-assessment scores*

ANCOVAs were conducted to determine if there were significant differences in delayed-test QSB scores across the four study strategy conditions. Examination of possible violations of assumptions for the ANCOVAs conducted in this section and adjusted mean scores are presented in [Appendix I](#). ANCOVAs conducted with overall delayed-test scores for QSB, while controlling for pre-test scores, did not reveal statistically significant differences across groups (Figure 13A). ANCOVAs of SPF sub-scores also did not reveal statistically significant differences across groups, but the SE group had the lowest mean sub-scores for structures and pathways of the four groups (Figure 13B).

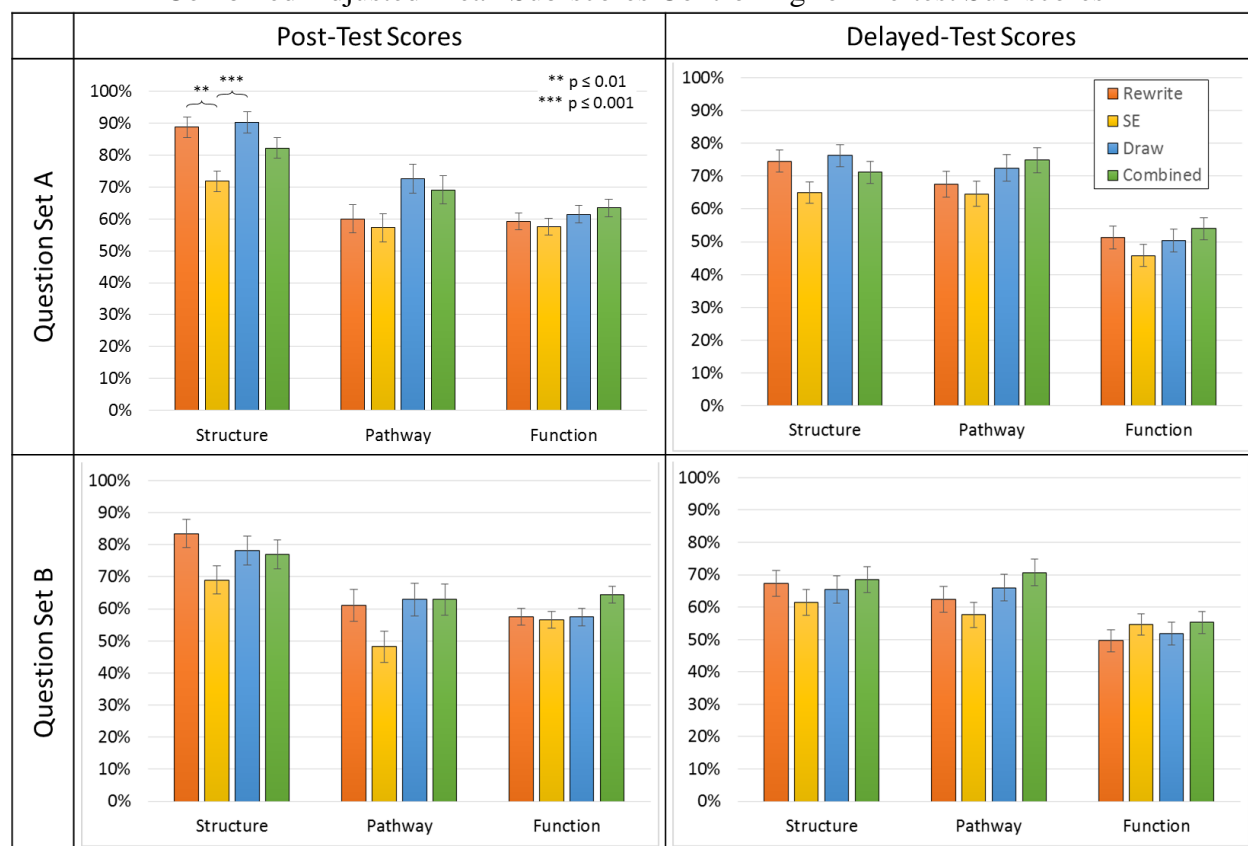


**Figure 13:** Adjusted mean delayed-test QSB scores with standard errors. **B.** Adjusted mean delayed-test QSB SPF sub-scores with standard errors.

f. *Summary of free-response results*

Pre-test scores (overall scores and sub-scores) were not statistically different across groups. However, given that small differences did exist, pre-test scores were used as covariates in subsequent ANCOVAs to control for them. Figures 4B, 5B, 6B, and 7B are presented side-by-side in Figure 14 for comparison.

Combined Adjusted Mean Sub-scores Controlling for Pre-test Sub-scores



**Figure 14:** Combined bar graphs for SPF sub-scores for the post-test (Figures 4B and 6B) and delayed-test (Figures 5B and 7B), by question set. These means were adjusted to control for pre-test sub-scores.

Mean SPF sub-scores for the SE group were the lowest of the four groups on both the post and delayed tests, for both question sets, except for delayed-test QSB functions sub-scores. Differences in group means for post-test QSA structures sub-scores were statistically significant, while some other differences for structures and pathways sub-scores were quite pronounced but not statistically significant. Post-test QSA and QSB sub-scores showed nearly identical patterns, but the magnitudes of the differences between groups on QSB were smaller. Delayed-test QSA and QSB sub-scores were also similar, except for delayed test functions sub-scores.

Overall, the patterns that seemed to arise from free-response sub-scores were as follows: 1) Participants in the SE group tended to have lower structures sub-scores than all of the other study strategy conditions. 2) Participants in the SE group also tended to have lower pathways sub-scores than other conditions, but sometimes less different from the Rewrite group. And 3) functions sub-scores tended to be similar across the four groups, with the SE group often having a slightly smaller mean sub-scores.

### 4.3 Figure Interpretation Task

The final task that participants were asked to complete, only at the end of their delayed-assessment, was the figure interpretation task. This task was presented after the drawing task and after answering the free-response questions. For the first time, participants were shown an

anatomical diagram of the CVS (Figure 3). They were asked to provide 18 labels for the components indicated and draw the direction of blood flow.

#### 4.3.1 Labels

The 18 components that were to be labeled on the figure are listed in Table 8 below. Student responses were scored according to how accurate and detailed their descriptions of these components were. For example, for component number 10 on the diagram, which pointed to the right atrium,<sup>4</sup> responses were scored with half credit for mentioning that the component was an atrium, even if they wrote “left atrium”. While the response “left atrium” means the student incorrectly determined which side of the heart it represented, he or she did understand that the top chambers of a heart are the atria. A similar partial credit scheme was used for valve components where half credit was assigned to responses that indicated that the component was a valve, but did not indicate the specific name of the valve.

**Table 8:** Labels for figure interpretation task that received full or partial credit.

	<b>Full credit labels (1)</b>	<b>Partial credit labels (0.5)</b>
<b>1</b>	Capillary	
<b>2</b>	Venules	
<b>3</b>	Vein	
<b>4</b>	Arterioles	
<b>5</b>	Artery	
<b>6</b>	Lung	
<b>7</b>	Aorta	
<b>8</b>	Pulmonary vein	
<b>9</b>	Pulmonary artery	
<b>10</b>	Right atrium	Atrium, left atrium
<b>11</b>	Tricuspid valve or Atrioventricular (AV) valve	Valve, bicuspid valve
<b>12</b>	Right Ventricle	Ventricle, left ventricle
<b>13</b>	Left Atrium	Atrium, right atrium
<b>14</b>	Bicuspid valve or Atrioventricular (AV) valve	Valve, tricuspid valve
<b>15</b>	Semi-lunar (SL) Valve	Valve
<b>16</b>	Left ventricle	Ventricle, right ventricle
<b>17</b>	Oxygenated blood	Deoxygenated blood
<b>18</b>	Deoxygenated blood	Oxygenated blood

All 74 participants who correctly described item 17 as oxygenated blood also described item 18 as deoxygenated blood. Of the 22 incorrect descriptions of these components, 5 students wrote oxygenated and deoxygenated blood but in the incorrect order, while the rest did not describe the state of blood at all (e.g., labeling these components as systemic and pulmonary

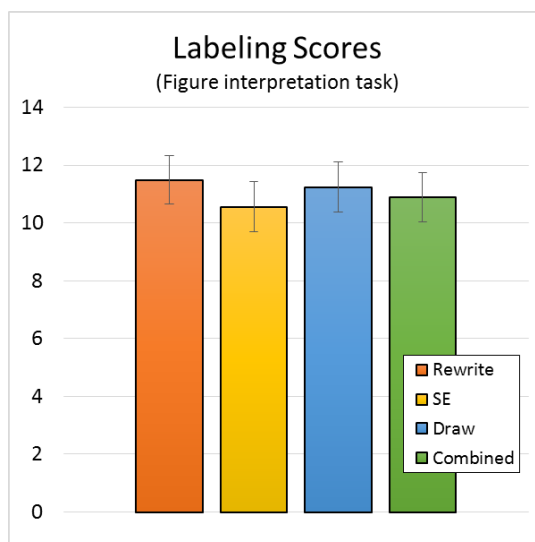
<sup>4</sup> If one were to trace the path from this atrium to the ventricle underneath it and out of the heart, one would find that that this side of the heart pumps to the lungs. Thus, it must be the right atrium. Many diagrams of the human body are oriented so that the body faces the reader (as if the reader were looking at someone face-to-face). Thus, diagrams are often drawn with the body’s left on the right side of the page and the body’s right on the left side of the page.

circulation, or veins and arteries). Thus, to prevent an inflation of student's scores by including points for both components 17 and 18, responses for these two components were combined and assigned one code: 1 if the two responses were correct and in the correct order, 0.5 if the responses were correct but in the incorrect order, and 0 otherwise.

**Table 9:** Mean scores (and standard deviations) for labeling components on figure interpretation task

Condition	n	Mean
<b>Rewrite</b>	24	11.48 (4.10)
<b>SE</b>	25	10.56 (4.32)
<b>Draw</b>	23	11.24 (4.19)
<b>Combined</b>	24	10.90 (4.19)

Average labeling scores, out of a maximum of 17, are presented by group in Table 9. A one-way ANOVA was used to determine if mean delayed-test labeling scores were statistically different across the four groups. There were no outliers as assessed visually by boxplots and data were somewhat normally distributed by condition. There was homogeneity of variances, as assessed by Levene's test of homogeneity of variances ( $p = 0.940$ ). The ANOVA determined that differences in means across groups were not statistically significant (Figure 15).

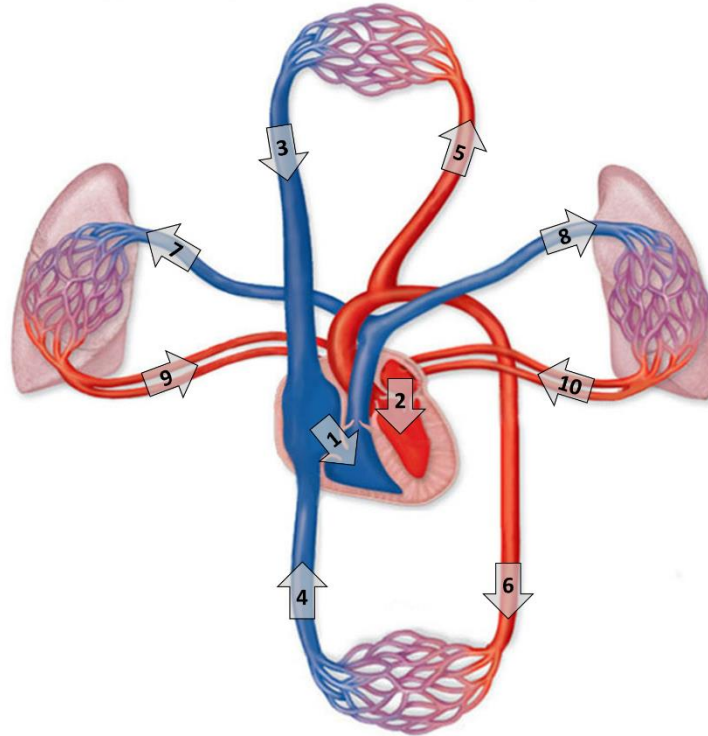


**Figure 15:** Mean figure interpretation task labeling scores with standard errors

#### 4.3.2 The Direction of Blood Flow

Participants were also asked to indicate the direction of blood flow (on the same figure they had used to label the 18 items) by drawing arrows onto the figure where they deemed appropriate. The number of arrows and their positioning around the figure varied from student to student. In order to analyze these various combinations of arrows, a minimum ten positions around the figure were identified where a direction needed to be indicated to demonstrate a

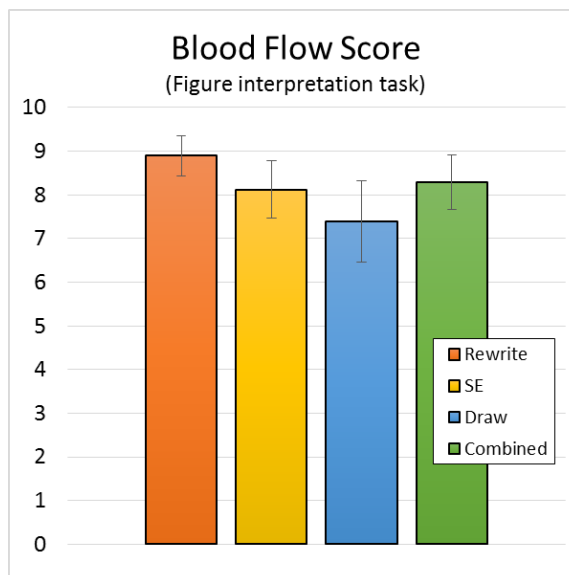
complete understanding of the direction of blood flow through the CVS. If a participant did not draw an arrow directly in any of these ten positions, the direction was interpreted by the two most adjacent arrows around that position. The ten positions were: 1) from the right atrium to the right ventricle, and 2) from the left atrium to the left ventricle; into the right atrium 3) from the upper body and 4) from the lower body; out of the left ventricle 5) to the upper body and 6) to the lower body; out of the right ventricle 7) to the right lung and 8) to the left lung; and into the left atrium 9) from the right lung and 10) from the left lung (Figure 16).



**Figure 16:** Ten positions indicating the correct direction of blood flow through the CVS.

A majority of participants in each group (67%-72%) were able to indicate the correct direction of blood flow at all ten positions on the figure. While not statistically significant, there were noticeable differences in mean blood flow scores across groups (Figure 17). Surprisingly, participants in the Draw group had the lowest mean score on this task compared to the other conditions. Given that so many participants were successful at this task, it was more informative to look at the range of incorrect answers rather than just the overall scores.





**Figure 17:** Mean blood flow scores with standard errors.

Participants who drew the complete path of blood flow correctly were likely to understand how blood flows through the CVS better than the other participants. However, it was difficult to draw conclusions about participants who drew the flow incorrectly, relative to each other. For example, a student who drew the entire circulation backwards may or may not understand the flow of blood better than one who only drew one sub-system loop backwards. The former may recognize the continuous cycle that is permitted by the structure of the pulmonary and systemic sub-systems but interpreted the image of the four heart chambers incorrectly, leading him or her to draw the entire system in the opposite direction. Such a drawing would be scored zero out of ten using our scheme. Meanwhile a drawing with only one sub-system drawn incorrectly might indicate that the student did not understand the continuous cycle of blood flow throughout the system, but could be scored as a six out of ten.

Given the complexity associated with comparing non-perfect blood flow scores, incorrect blood flow drawings were further analyzed by grouping them by the types of mistakes made (Table 10). A figure with blood drawn flowing consistently in the opposite direction in the ten places indicated was categorized as “Backwards.” A figure with one incorrectly drawn loop (to either lung, or to either the upper or lower body) was categorized as “One wrong loop.” A figure with both loops in either the pulmonary or systemic circulation drawn incorrectly was categorized as “One wrong sub-system.” And a figure with more than one sub-system drawn incorrectly was categorized as “Other.”

**Table 10:** Paths of blood flow categorized as correct or by type of error.

Condition	n	Correct	Incorrect			
			Backwards	One wrong loop	One wrong sub-system	Other
<b>Rewrite</b>	24	18	0	2	2	2
<b>SE</b>	25	18	1	1	1	4
<b>Draw</b>	22	16	2	0	0	4
<b>Combined</b>	23	16	1	2	3	1

Either six or seven participants in each group did not draw the path of blood flow through the 10 positions correctly. Interestingly, of the participants in the Draw group who had imperfect blood flow scores, they either drew the path completely backwards, or very incorrectly with no discernable pattern. Whereas, participants in the other groups ranged in the number of incorrect loops they indicated in their systems.

Participants, especially in the Draw group, may have been confronted with a conflict between their models of the CVS as they had been drawing it and the anatomical figure presented to them on this task. When participants drew the CVS, most drew blood coming out of ventricles in a downward direction from the bottom of the heart. Thus, students may have had trouble recognizing that the anatomical figure depicted blood pumping upwards out of the heart. In the figure, the blood vessel that would most closely resemble the ones students draw leaving the heart (in a downward direction) would actually be the inferior vena cava, where arrow 4 is located in Figure 16. When the directions indicated at position 4 were counted for participants with incorrectly drawn blood paths, all but one participant in the Draw group indicated that blood flowed downward in this blood vessel (Table 11).

**Table 11:** Summary of blood flow direction indicated through position 4 of participants with incorrectly drawn blood paths.

Condition	n	Position 4	
		Upwards*	Downwards
<b>Rewrite</b>	6	4	2
<b>SE</b>	7	4	3
<b>Draw</b>	6	1	5
<b>Combined</b>	7	4	3

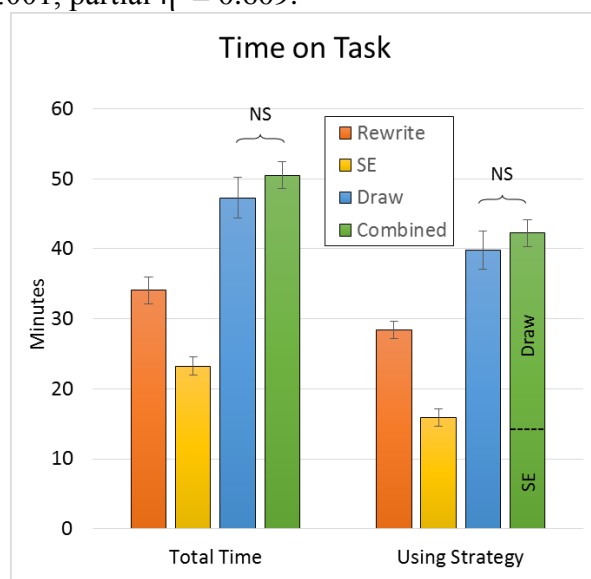
\*The correct direction

In summary, a majority of participants in each group were able to correctly indicate the direction of blood flow through 10 specific positions on an anatomical figure of the CVS. Of the participants who were unsuccessful, it was inaccurate to use their blood flow scores to make conclusions about their understanding relative to one another. To enhance this analysis, incorrectly drawn blood flow paths were categorized and we found that participants in the Draw group tended to draw their paths either completely backwards, or very incorrectly. These participants were also more likely to indicate that blood flows downward out of the heart at position 4.

## 4.4 Time on Task

Time on task was not a measurement on the formal assessment itself, but it could be used to investigate if and how differences in these times were related to performance on the assessments. This section shows the results of different categories of time being measured during the independent study sessions and how these measures were related to performance on the formal assessments.

Using the drawing strategy to study the CVS text was expected to require more time than the SE strategy, and the Combined strategy would take even more time. Mean times (in minutes) spent studying the material in total and spent perceptibly using a study strategy are shown in Figure 18. The SE group spent the least amount of time overall and perceptibly using their assigned study strategy. The Rewrite group spent significantly more time on both of these than the SE group, and the Draw and Combined groups, significantly more than the Rewrite group. A four-group, one-way ANOVA determined that differences in total time spent studying the CVS text were significantly different across groups,  $F(3,83) = 39.228, p \leq 0.001$ , partial  $\eta^2 = 0.586$ . Specifically, each group differed from one another at the  $\alpha=0.01$  level, except between the Draw and Combined groups. The same was true with time spent perceptibly using the study strategy,  $F(3,83) = 43.176, p \leq 0.001$ , partial  $\eta^2 = 0.609$ .



**Figure 18:** Mean times spent studying the material in total and perceptibly using a study strategy, and standard errors. All post hoc comparisons between pairs of groups were significant at the  $\alpha=0.01$  level except where indicated as not significant (NS).

It is important to note that the times spent using strategies were not specific for how the strategy was being used or if it was being used correctly. Specifically, for participants in the SE group, time spent using the study strategy was measured as time that the learner spent talking aloud. While many of their utterances are likely to not qualify as proper self-explanations, these times were included in these rough measurements. The quality of these self-explanations generated during independent study sessions will be examined in Part 2 of this dissertation.

Correlation analyses were conducted to see if there were any relationships between the amount of time a participant spent using their study strategy and their performance on the post- and delayed-tests. The analyses were conducted using times participants spent using their study strategies by group and in pairs for common strategies. For example, correlations were tested

with assessment scores against times from only the SE group, and also against combined times from the SE group and time spent self-explaining in the Combined group. Formal assessment scores were condensed by averaging scores on QSA and QSB and on both the post and delayed tests. Short-term and long-term DMM gains were also averaged. (A full correlation matrix with individual scores is presented in [Appendix J.](#))

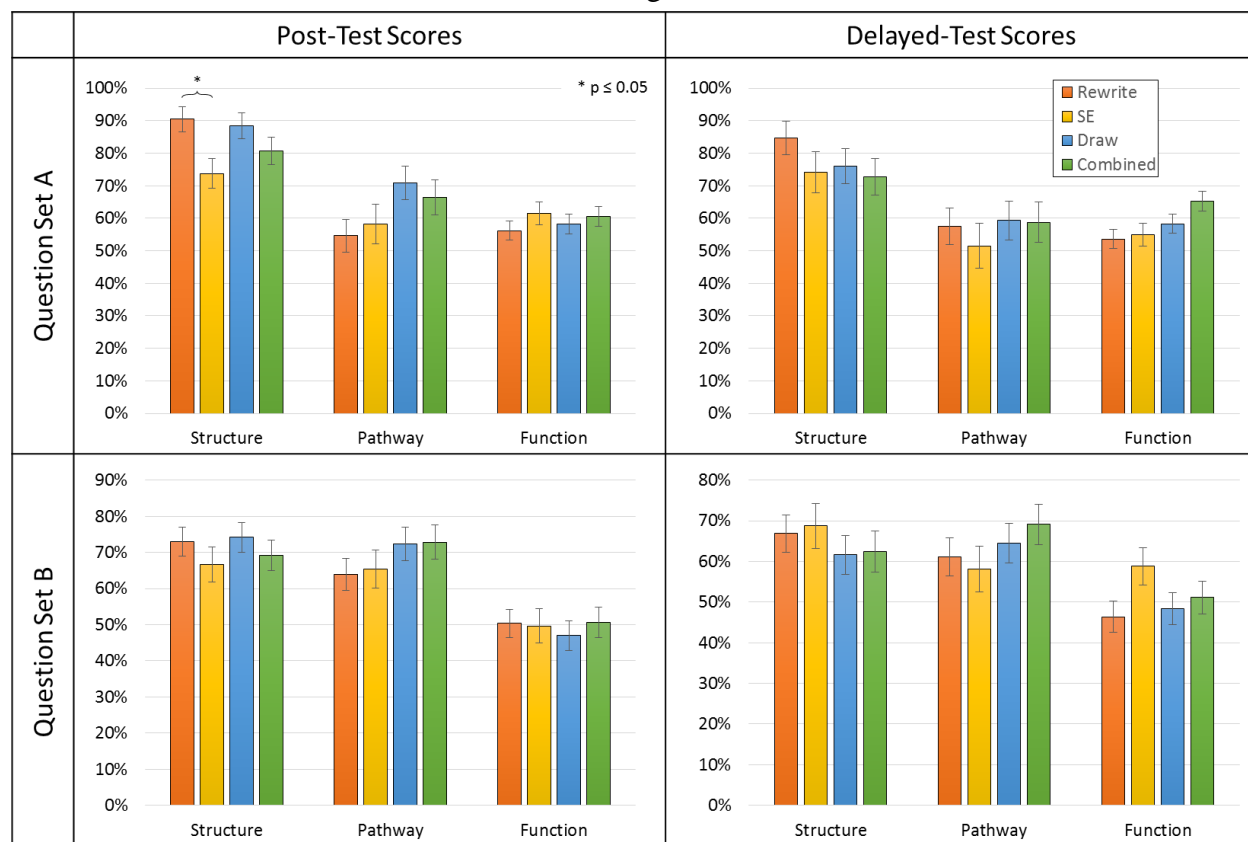
The only statistically significant correlation found between time spent using study strategies and assessment scores was with overall times spent drawing and DMM scores ( $r = 0.33$ ,  $n = 43$ ,  $p < 0.05$ ; Table 12). Other correlations were positive and moderate in strength, but not statistically significant. Time spent using study strategies in the SE group and the Draw group were weakly to moderately correlated with some formal assessment scores. Overall, the correlations between times spent using study strategies and formal assessment scores were not strong.

**Table 12:** Correlation matrix between times using study strategies and formal assessment scores

Scores	Time using Strategy by Condition				Collective times	
	Rewrite (n=19)	SE (n=24)	Draw (n=20)	Combined (n=23)	SE time (n=47)	Draw time (n=43)
<b>Total S</b>	-0.13	0.17	0.39	-0.18	0.12	0.03
<b>Total P</b>	-0.07	0.24	-0.22	0.13	0.24	-0.11
<b>Total F</b>	-0.16	0.20	0.23	-0.02	0.22	0.01
<b>DMM</b>	-0.15	0.19	0.36	0.11	-0.01	<b>0.33*</b>

Given that the SE group spent significantly less time studying the material than the other groups, an analysis was conducted to see how sub-scores would differ if time was controlled for. ANCOVA tests were conducted using pre-test sub-scores *and* time spent using a study strategy as covariates. Adjusted mean scores for post and delayed tests for both question sets are presented in Figure 19.

### Adjusted Mean Sub-scores Controlling for Pre-test Sub-scores and Time Spent Using Study Strategies



**Figure 19:** Combined bar graphs for SPF sub-scores for the post-test and delayed-test, by question set. These means (and standard errors) were adjusted to control for pre-test sub-scores and time spent using study strategies.

Many of the patterns across groups for each adjusted sub-score persisted after controlling for time (i.e., compared to sub-scores in Figure 14, where time was not controlled for). The difference in structure sub-scores between the Rewrite group and SE group on Post-test QSA remained significant ( $p < 0.05$ ). In general, most of the adjusted mean SPF sub-scores for the SE group increased when time was controlled for, resulting in the mean sometimes exceeding the mean sub-score for the Rewrite group but rarely exceeding the Draw and Combined groups'. The SE group still had lower mean sub-scores than the Draw and Combined groups for nearly all structures and pathways sub-scores.

## 4.5 Discussion of Formal Assessment Results

The Rewrite group was originally meant to serve as a control group by having participants copy the CVS text verbatim, assuming they would not use other strategies such as self-explaining or drawing. However, many of these participants spent time with each passage in silence, presumably using some sort of strategy to study the material. Additionally, it is likely that participants were able to simultaneously rewrite the text and use another strategy to learn as they wrote. Thus, rather than use the Rewrite group as a “control,” it is more useful to consider it another experimental group and make comparisons most often to the SE group since self-explanation has been repeatedly shown in prior research to be an effective study strategy for learning from text.

While the differences in most of the post and delayed test SPF sub-scores between groups were not statistically significant, the patterns in the data are still meaningful and are used here to make claims about the effects of the different study strategies. What follows is a discussion of each study strategy condition, considering theoretical explanations for each.

#### 4.5.1 Rewriting

---

Of the four study strategy conditions in this project, I hypothesized that participants in the Rewrite group would learn the least about the CVS because copying text would have been the least conducive to processing information, compared to self-explaining and drawing. However, participants in the Rewrite group did more than just copy text during their independent study sessions, and this was reflected in their video data as well as their assessment scores. On average, participants in the Rewrite group showed higher learning gains about structures and pathways, and also higher DMM gain scores than participants in the SE group. Because the processing that Rewrite students did was largely unobservable, there are many possible explanations for these differences.

What is obvious is that, during their independent study sessions, it took most students in the Rewrite group more time to hand-copy paragraphs of text than students in the SE group to self-explain, as seen when comparing times on task. The rewriting strategy was meant to serve as a task that took more time to complete than re-reading, the condition used as the control group in Chi et al. (1994), and perhaps as much time as self-explaining. In the end, however, it required significantly more time than self-explaining. I assumed that rewriting text would still have a high extraneous cognitive load, reducing germane load (Sweller, Van Merriënboer, & Paas, 1998). In other words, the physical act of copying text would require some mental capacity, taking away from what could be used to more efficiently learn the material. Therefore, I presumed that because a participant in the Rewrite group spent more time with the material than one in the SE group, did not mean that he or she would learn more because this time was not being used for efficient learning (relative to self-explaining).

The fact that participants in the Rewrite group learned more about the CVS on multiple measures than those in the SE group immediately directs attention to the quality of their independent study sessions. What were participants in the Rewrite condition doing while studying, other than copying text? It may be that the extraneous cognitive load for rewriting that was hypothesized to detract from germane load was less strenuous than anticipated. As a result, participants in the Rewrite group were able to use their time with the CVS material – simultaneously while rewriting and/or during silent times – to learn effectively via the execution of other study strategies. It is possible that these participants self-explained covertly, re-read the passage, visually imagined the system, had more time to recognize gaps in their understanding, any combination of these strategies, or something entirely different. Regardless of the method(s) used, on average, these students used silent study strategies that were more effective than self-explaining aloud.

The fact that participants in the Rewrite group took longer to use their study strategies with the CVS text than those in the SE group also directs our attention to what impact the quantity of the time spent with the material could have on each group. It is possible that the Rewrite strategy forced participants to slow their studying of the material to the pace of their writing, allowing them to spend more time processing the information, generating inferences about the material, reorganizing their knowledge structures, or using any other process that promotes learning. The self-explanation strategy did not have the same slowing effect. In fact,

even after training and practice, many participants could read multiple lines in the text aloud and then not know what to self-explain. Sometimes participants who were self-explaining would verbally express that the text was clear and there was nothing that needed to be explained, and would simply move on.

On another note about pace, participants in the Rewrite group were more likely to read the passages twice, once for general comprehension and once again as they wrote down every word. In the SE group, on the other hand, many participants self-explained *as* they read through the passages, rarely reading through a first time for general comprehension. Reading through the material twice may have allowed participants in the Rewrite group opportunities to catch information that they missed while reading the first time, or correct ideas that they misinterpreted the first time.

Given the lack of data on what *else* Rewrite students were doing during their independent study sessions, there are many explanations that this study is not capable of examining for why they tended to outperform participants in the SE group. I believe that the Rewrite condition allowed students to enact other, more effective, silent learning strategies. While these data cannot provide us answers to what processes students in the Rewrite group were using, they have shown us that asking students to use the rewriting strategy to study text was better for student learning than asking them to self-explain.

#### 4.5.2 Self-explaining

---

It was unanticipated how poorly participants in the self-explanation group performed on the formal assessment measures in comparison to the other conditions. It is expected that if this study had a control group with participants who simply read the CVS text aloud twice, as in Chi (1994), participants in the SE group would outperform that group. This is reasonable considering that the CVS text and assessment questions were mostly identical to those used in Ainsworth & Loizou (2003) and Chi (1994), and that self-explanation has been shown rather consistently in the literature to have a positive effect on learning (e.g., Chi et al., 1989; Chi, 1994; Chi & VanLehn, 1991; McNamara, 2004). Furthermore, rather than only consider these results to mean that self-explanation was the least effective of the four strategies, it is also productive to frame these results within the idea that self-explanation was effective and the other three study strategies were *more* effective.

The finding that the SE group did not perform as well as the other groups on measures of learning about structures and pathways but did on learning about functions speaks to the type of learning that may and may not be supported by self-explaining. During their training sessions, participants in the SE group were taught to make connections among the passages and sentences they were reading and with their prior knowledge. They were encouraged to elaborate on ideas and make inferences about the content. These training sessions did not specifically direct participants to focus on components of function over other categories, but (as will be discussed in more detail in Part 2) many of the self-explanations that participants generated during their independent study sessions were about how things in the CVS worked (and less about how things were structured or where blood flowed).

While the results presented thus far do not allow for the microgenetic analyses required to make strong claims about the learning mechanisms behind the self-explanation effect, sufficient trends exist to examine the results through the lenses of mechanisms proposed in prior research. These mechanisms for how self-explanation promotes learning are: 1) through generating inferences and 2) through revising mental models (Chi et al., 1989; Chi et al., 1994; Chi, 2000).

This discussion also go a step further to consider the extent to which each of these mechanisms plays a role in the learning of structures, pathways, and functions within the CVS.

*a. Generating Inferences*

Generating inferences while reading text is inversely related to how coherently the text is written. That is to say, the *less* coherent the text is, the more self-explanations a learner tends to generate while reading (Ainsworth & Burcham, 2007), some of them qualifying as inferences. To generate such inferences requires that a reader conclude information that is not explicitly stated in the text, by integrating information across different sentences or with prior knowledge, or by implying information based on the meaning of words (Chi, 2000). In the context of learning about the CVS, participants may feel less inclined to generate self-explanations about structures and pathways because this type of information may seem more straightforward and be understood directly from the text, without needing to make inferences. For example, the following sentences from Passage 2 seemed simple to many participants: Each lower chamber is called a ventricle. Each upper chamber is called an atrium. Many participants did not follow reading this text about heart structures with self-explanations about structures. Likewise, texts about pathways, such as Blood from the right ventricle flows through the semilunar valve into the pulmonary artery and then to the lungs. also appear straightforward and did not stimulate many learners to generate self-explanations about pathways. In general, in the CVS text, structures are usually named and described, and blood is stated to flow from one organ to another. These lines in the CVS text generally did not stimulate the generation of verbal inferences.

On the other hand, the CVS text contained many sentences that required generating inferences about functions, nearly prompting learners to generate explanations. For example, reading a sentence about a general or abstract concept, such as Diffusion is the process by which molecules spread from an area of greater concentration (or density) to an area of lesser concentration. is likely to compel participants to generate inferences about how the concept applies to the CVS. Applying concepts in this way, usually explained *how* or *why* something in the CVS worked a certain way. In this example, participants would often infer that the process of diffusion applied to the CVS in the capillaries where the concentration gradients of oxygen and carbon dioxide caused diffusion of these molecules across the blood vessel's membrane.

I presume that students asked to read the CVS text and self-explain are more focused on answering "why" and "how" questions for themselves, rather than "what" and "where" questions that often do not motivate explanation. The former might be considered deeper types of questions because they may require more coordination of knowledge and inferences to answer, for example forming causal links to explain the purpose of various components in the CVS. Whereas the latter types of questions might be answered by referring directly to a line or two in the CVS text. Thus, if participants are generating more self-explanations about functions, they may be learning more, and more deeply, about this knowledge category than structures and pathways. This makes sense given that participants in the SE group in the study did not fall behind the other study conditions on learning measures about functions, but they did on structures and pathways. In sum, the SE condition appeared to help participants learn about functions in the CVS by encouraging the generation of inferences about functions but perhaps did so to a lesser extent for structures and pathways.



*b. Revising Mental Models*

The second mechanism for how self-explanation has been proposed to promote learning is through the revision of mental models. Self-explanations are sometimes stimulated by a learner's recognition of conflict between their mental model and information presented in the text, prompting him or her to consider a revision (Chi et al., 1994). This recognition is, at least in part, determined by the extent or strength of the conflict. A conflict that has very little effect on one's mental model may go unnoticed, whereas a conflict that could affect multiple parts of the system would be difficult to ignore.

One explanation for how revising mental models might work considers when there is a gap in one's original mental model. If the information about structures and pathways does not conflict with the learner's mental model because that part of their model does not yet exist, new information presented through text in a coherent manner might simply be integrated. That is to say, the information simply fills the gap in their mental model. As stated in Chi (2000), "sentences containing information that fill gaps of missing knowledge can simply be assimilated without much self explaining" (p.198). This is consistent with my claims that participants in the SE group tend to self-explain about functions more than structures and pathways. Self-explanations may not be necessary when knowledge about structures and pathways can simply be assimilated. Furthermore, while this assimilation assumes that the new information about structures or pathways was learned, the extent with which this information was learned may be superficial, given that the "conflict" was relatively superficial. This is in contrast to deeper conflicts which, when resolved, may lead to deeper learning.

With a goal of deep learning, it may be more important for us to understand conflicts that are not due to gaps in a learner's mental model but differences between what exists in a learner's model and the text's model. While these conflicts may generally be stronger than those involving knowledge gaps, they will range in the extent to which the conflicting knowledge is embedded within the learner's mental model, and thus individually range in their strength of conflict. As mentioned above about textual coherence, information about structures and pathways may be presented in the text in a more direct way than functions, perhaps making conflicts appear weaker, less noticeable, or less frequent. Conflicting information about what a structure is called or how it is built might easily be "replaced." Similarly, conflicting information about where or in which direction blood flows might be something that can also be replaced. Internal knowledge structures about CVS structures and pathways may not be as deeply embedded in and connected with other information about the CVS as functions.

A conflict between functions as understood in one's mental model and as presented in the text is likely to require revisions of knowledge structures deeply embedded within other knowledge structures. Chi (2000) refers to similar conflicts as violations:

a violation requires a recognition that a piece of knowledge is violated in some *causal way*, in the sense that this flawed belief has some implication for additional consequences, so that a repair is not merely the case of replacing an isolated incorrect belief with the correct one. (p.200)

A violation is more likely to be about a function than a structure or pathway. This is because knowledge that describes how a component in the CVS works should lie in a network of knowledge of other components in the CVS (including structures and pathways) that considers its effects on and relationships to those other components. For example, if a learner's mental model can be described as a Single Loop with Lungs and he or she comes across text that describes two loops, the models conflict drastically. In order to revise the learner's model, the

pathway of the single loop between the lungs and the body must be broken, and the blood flow must be redirected back to the heart. While these may seem to be revisions of pathways, the deeply embedded knowledge that is the foundation of understanding why this loop revision is necessary is the function of the heart. If a learner realizes that the function of a mammal's heart is to efficiently pump blood to every cell in the body, distributing oxygen and collecting waste, he or she may also realize that a double loop model: 1) allows sufficient force for blood to be pumped to the lungs and separately to the whole body, 2) allows blood flow to slow in capillaries equally in the lungs and at body cells, permitting efficient gas exchange, and 3) explains why a septum and four (rather than two) chambers exist in the heart. In this example, we see that functions are deeply embedded in an understanding of many other parts of the CVS, including structures and pathways. Thus, violations occur when there is conflict between new information and deeply embedded knowledge within a mental model, and this knowledge is seemingly more likely to be about functions than structures or pathways. Furthermore, if a self-explanation provides a learner with the opportunity to discover violations and the learner resolves them by revising their mental model, this may result in deep learning.

What has just been discussed about how the self-explanation strategy encourages learners to revise their mental models may seem contradictory to the results about the DMM scores, given that the SE group showed the smallest gain in these scores. However, keep in mind that while this gain was the smallest, it was still significant. So, it is possible that self-explaining is effective for improving DMM scores, but the other conditions are more effective. It is also possible that the mismatch between the mode of studying (reading text and verbalizing explanations) and this mode of testing (drawing) put participants in the SE group at a disadvantage. Finally, it is possible that DMM scores are just not good measures of learners' complete mental models. Drawings are not sensitive tools for examining student understanding of functions since functions are difficult to illustrate. Meanwhile, structures and pathways can be easily included in drawings, and in this task, pathways were explicitly asked for and structures were necessary for context. Functions, on the other hand, were not particularly necessary to complete a drawing.

#### 4.5.3 Drawing

---

While participants in the Draw group learned as much about functions of the CVS as participants in the SE group, they were also able to learn more about structures and pathways. The Draw group was significantly better at improving their drawing mental models in this study. Surprisingly, where the Draw group did not meet expectations was on the figure interpretation task, where they performed slightly worse than the other groups.

It is fair to assume that there are some mechanisms for learning that are shared between self-explaining and drawing, such as promoting the generation of inferences. However, the two study strategies are so different from each other that discussing both in terms of the same underlying mechanisms takes for granted the disparate benefits of each. Instead, to highlight the process of drawing to learn, this section will discuss the results of the Draw group with only some reference to the SE group for context. Then, the section will conclude with a direct comparison of the two.

Participants in the Draw group displayed significant learning of the three SPF knowledge categories. I believe that the drawing strategy helped students to learn about structures and pathways in a way that self-explaining did not. And while the Draw group did not differ from the

SE group in how much they learned about functions, participants in the Draw group may have learned about functions in a different way, and possibly learned a different set of functions.

*a. The Nature of Drawing the CVS*

In the context of this study, drawing inherently places an emphasis on structures and pathways. Participants drew heart chambers, lungs, other body parts, and the blood vessels connecting them as the text presented the scientific model of the CVS. As these components were named and described in the text, participants were forced to give them shapes and spatially arrange them in their drawings. They could focus on structures first and then draw the pathways of blood that connected them, or focus on pathways and include the structure of their origins or destinations. Either way, asking learners to draw as they read about the CVS makes features of structures and pathways explicit. Learners were required to face conflicts and uncertainties they had about structures and pathways in the text in order to include the information in their drawings.

Van Meter & Garner (2005) presents a detailed theory about the cognitive processes that underlie drawing to learn from text. It proposes that learning involves selecting, organizing, and integrating verbal and nonverbal representations of knowledge, which are distinct from each other. Applying their theory to this study gives us the following possibility. From the CVS text, learners select key components about structures, pathways, and functions to organize and form internal verbal representations of the information. Verbal stored knowledge is then activated to facilitate comprehension. Once the learner begins considering options for drawing, nonverbal stored knowledge is activated. At this stage, it is more likely for a learner to have stored nonverbal (i.e., visual) knowledge of structures and pathways (e.g., what the heart and blood vessel loops look like), than of functions (e.g., why atria pump first and ventricles pump second). As such, integration between the verbal and nonverbal representations of the CVS favors structures and pathways.

Even in deliberately creating drawings about functions, structures and pathways are unavoidably integrated. In order to describe how or why something works, the “thing” must first be defined. Thus, in any attempt to represent the function of the heart in a drawing, the heart structure must first be outlined. Likewise, in order to explain how or why blood must travel to the lungs to acquire oxygen, blood should be drawn coming from somewhere and going to the lungs, and then leaving the lungs and going somewhere else. In essence, if a learner wants to represent function in their drawing, they are (consciously or subconsciously) also representing structures and/or pathways.

The previous examples explain how drawing the CVS places strong direct and indirect emphases on structures and pathways, but I believe that the benefit drawing has for deeper processing is actually what allowed participants in the Draw group to learn about functions. Considering that functions are about how and why components work, there are sometimes issues of animation or the connection between structure and function that cannot be drawn as easily as structures and pathways. Accordingly, it is likely that learners need to strategize how they can represent a function in their drawing. This process might lead students to consider aspects of the function or connections between structures and functions that they otherwise would not consider (Van Meter et al., 2006). For example, when considering diffusion across a membrane, mentally a student may only consider molecules moving in one direction, from an area of high concentration to low. But when forced to draw diffusion, the student may realize that he or she needs to consider a representation of the state that exists after diffusion is complete (at

equilibrium), or that molecules can also be moving in the other direction. The effort associated with generating a representation that appropriately encompasses the dynamics of the system or other complex relationships can lead to deeper learning of the content (diSessa et al., 1991). Therefore, while it may be more challenging for a student to represent function in a drawing, this challenge may be what allows the student to learn that information more deeply.

*b. Computational Offloading*

Less specific to drawing the CVS, computational offloading can be used to explain how drawing, in general, can support learning. A learner is able to distribute the cognitive effort required to process all of the information onto an external representation. Again, the CVS is complex and while the CVS text is well-written and widely used, it is still difficult and dense with information and terms. Drawing permits learners to offload some of this information and focus on smaller, more manageable subsets or chunks (Bransford, Brown, & Cocking, 1999). As a result, the chances of generating an incorrect inference may be reduced and perhaps deeper processing of each chunk can be accomplished.

Since they cannot include all possible details of the CVS in their drawings, learners may only include information that they deem important and representable. This means that drawings will put graphical constraints around the possible inferences that the learner can make (Scaife & Rogers, 1996). Spatial constraining occurs when the learner makes decisions about the relative sizes and arrangement of CVS structures. Furthermore, temporal constraining is inherent in the static nature of drawings. For example, in the case of illustrating the pumping of the heart, the learner must consider a representation of the heart in its contracted state and another representation of the heart in its relaxed state. Attempts to combine the two into one representation might produce a confusing drawing. With these constraints in place, the drawing that results will contain inferences that the learner made about how the information looks and how it is all put together. The learner can only make a restricted set of inferences based on the information that is included in the drawing.

The inherently simplified version of the CVS in a learner's drawing puts a focus on the loops through which blood flows throughout the system. Students in this group, essentially, had practice generating drawings that explained how blood traveled through the system before generating their drawings on the formal assessments. In combination with the attention to including structures and pathways in ones drawing, this explains why participants in the Draw group performed so well on drawing mental model tasks. As previously mentioned, the DMM measure focused more on structures and pathways because functions are difficult to include in and interpret from a drawing. Moreover, the drawing task essentially asked Draw students to generate a representation similar to those they had been generating throughout their study sessions.

Once structures and pathways are established in a drawing, the learning strategy may shift from learning by *generating* a representation to learning by *using* one's representation. While they, hopefully, include the same important pieces of information as the text, learners' drawings are likely to be computationally more efficient due to the constraints discussed above. Rather than searching the text or one's memory to recover related information, a drawing can put all related information in the same place. Saving a learner from this effort lets him or her focus on other aspects of the system, such as functions.

Seeing all of the components of the CVS drawn together in one image can trigger learners to consider the relationship between these components. For instance, why do arteries

need to branch into small capillaries? Or why can't blood be pumped from the lungs directly to body cells? A drawing that includes all components of the CVS begs the learner to form a coherent model of the system. By this, I mean the learner can mentally test their drawing model to simulate blood flowing through the system and components functioning or not functioning as they are meant to. The septum is an example of a component that has a function which becomes very apparent after considering how all parts of the CVS come together. As one follows the state of blood as it passes through the left and right sides of the heart, it becomes clear that the septum must be a solid divider in the heart to prevent oxygenated and deoxygenated blood on each side from mixing. Without the cognitive load of structures and pathways on a learner's working memory, ideas about the relationships and roles of components in the CVS can be brought forward.

Overall, it seems that the nature of the CVS requires that structures and pathways be explicitly included in a learner's drawing, and once those components are drawn, cognitive effort can be directed towards understanding functions. Learning in all three SPF knowledge categories are supported by drawing. That being said, it was not surprising that the Draw group performed very well on all post- and delayed-test SPF measures, including the drawing task. The only surprising result from the formal assessment was participants' performance on the figure interpretation task.

One might assume that since the mode of studying for the Draw group was visual, they would perform well on a visual assessment. However, generating a representation is a different process from interpreting a representation generated by someone else. Given that participants in the Draw group had generated many representations of the CVS during their study sessions, they were likely to try to fit the image of what they had been drawing onto the figure presented to them. Perhaps, the repeated offloading of information about structures and pathways onto drawings made this process mechanical. As a result, students did not realize or even pay attention to the fact that the directionality in the figure does not match that in their drawings. This makes sense given that many students drew blood exiting the right ventricle in a downward direction and most of the incorrectly drawn pathways in the Draw group indicated that same downward direction at position 4 in Figure 16. In summary, while SPF knowledge categories are all supported by drawing, some cognitive offloading may result in less processing of information that becomes mechanically represented in a drawing.

### *c. Drawing versus Self-explaining*

For learning, the most obvious advantage of drawing over self-explaining is that the outputs are physically lasting. A learner can look at their drawing when finished and review the representation of the ideas he or she had considered since the beginning of the passage. In contrast, self-explanations are fleeting and no physical evidence of them remains after they are spoken. Thus, if a learner is to develop a coherent understanding of the information presented in a long passage through self-explaining, he or she must chunk or commit as much of this information to memory as possible. Considering the complexity and density of the information in the CVS text, it would be valuable to reduce this load on working memory as one reads the text. Under these circumstances, drawings can be utilized to reduce the load by offloading cognitive effort onto an external representation.

Self-explaining, on the other hand, also has a main advantage over drawing. Self-explanations can be explicitly as detailed and thorough as one wants them to be. If an idea can be expressed in words, it can be used in a self-explanation. In contrast, the accuracy and detail of a

learner's drawing is restricted, not only by the medium and the artistic abilities of the learner, but also how experienced the learner is with representations and representing. Generally, the only limit on the extent of a self-explanation is the explainer's vocabulary. Specific to the CVS, details and explanations about *functions* may be more easily expressed in self-explanations than drawings. Again, the issues of motion and purpose are not easily transmitted in a drawing.

The next factor to consider is time. As supported by the results, self-explaining can be done in a much faster manner than drawing. If this meant that self-explanation was a more efficient study strategy than drawing, this would be a clear advantage. However, it is possible that less time spent with the material is related to less learning. In the results, the SE group spent the least amount of time with the learning material, and performed the worst of the four groups on most measures. While this relationship was not proven to be causal, it is reasonable to assume that if a learning strategy caused a learner to slow down and spend more time with the CVS text, they would learn more.

I believe the extra time spent in the Draw group compared to the SE group include two factors that would enhance learning: 1) time spent recognizing confusion and conflicts and 2) time spent attempting to resolve these issues. Considering the explanations proposed earlier about how aspects of the CVS are made explicit through a drawing, learners will come to notice confusion or conflicts as they add more representations of the text to their drawing. This is less likely to happen with self-explaining. Even though a learner can go through each sentence of text and provide an explanation about the content of each, this does not mean that a learner will notice conflicts. One could very well provide an explanation for one sentence and an explanation contradictory to it immediately after for the next sentence and not notice the conflict (having forgotten the details of what he or she just said). With drawings, however, it would be much more noticeable to add a piece to a drawing that is inconsistent with what has already been drawn on the paper.

In addition to recognizing that these issues exist, addressing and resolving them may also be more likely to occur when using the drawing strategy than the self-explaining. With the self-explanation strategy, if a learner comes across a line of text that is confusing, he or she may vocalize what the confusion is about and then continue to the next sentence, or may simply ignore the conflict altogether. With drawings, on the other hand, I believe that learners are more likely to attempt to resolve a conflict in order to complete their drawings. In this study specifically, this effect was enhanced by the training sessions because participants in the SE condition were encouraged to ask questions and make predictions about what would come next in the text. In a way, it may have seemed to participants that having questions and being uncertain about what one was reading were sufficient outcomes of self-explaining.

Finally, let us consider some of the differences in cognitive processes between self-explaining and drawing. As presented earlier, self-explanation has been argued to promote learning through the generation of inferences and revision of mental models. Learning by drawing can also be explained via the same mechanisms. A transformation from text to drawing requires that a learner make inferences about the system to fill in what the text does not say about the details of size and arrangement. Moreover, in addition to changes that a learner makes to their drawing, every addition to their drawing and every iteration of their drawing for different passages are revisions of the learner's mental model. Such revisions are less obvious when self-explaining.

As well as generating inferences and revising one's mental model, drawing has added processes that self-explanation does not. In terms of Van Meter's Generative Theory of Drawing

Construction (Van Meter et al., 2006; Van Meter & Garner, 2005), the processes of selecting, organizing and integrating when learning from text and generating a drawing are very specific to transformations from text to drawings. Van Meter compares this theory for learning from text by drawing to learning from an illustration to drawing. While her comparison is of a different process, it is a situation where the mode of learning matches the output – both are visual. In the case of the SE group, the mode of learning was verbal and the output was verbal. In drawing from text, the relationship between the dual modes of verbal and visual representations is much more complex than with the single mode. Verbal and nonverbal representations must be linked and integrated in order for a drawing to be generated. This challenging process is likely to be what underlies learning from text by drawing.

This section highlighted many of the main features of drawing-to-learn and made several comparisons to the self-explanation strategy. Separately and in comparison to one another, the results from the SE and Draw groups provide great insight into possible differences that may exist between the two strategies for learning from text. A combination of the two strategies into one condition may present additional insight into how the two may complement or hinder each other. Additional comparisons and contrasts between drawing and self-explaining are presented in the next section where results from the Combined group are discussed.

#### 4.5.4 Combined – Self-Explaining and Drawing

---

Initial hypotheses about the Combined group assumed that the benefits of the self-explanation strategy and drawing strategy would be additive. That is, the participants in the Combined group would learn the most about the CVS compared to the other groups. However, this was only the case for measures of understanding about functions. The Combined group performed at levels between the SE and Draw groups on most free-response measures of structures and pathways, and on the drawing task (which, as mentioned earlier, is believed to be reflective of understanding of the structures and pathways). Their scores were similar to the SE group on the blood flow portion of the figure interpretation task.

The fact that the Combined group had the highest mean sub-scores for functions across the groups suggests that perhaps for this knowledge category, the benefits of the two study strategies are indeed additive. As mentioned when discussing the SE group, self-explanations are good for learning about functions because it makes more sense to explain ideas about how and why things work, than explain structures or pathways. Drawing is also useful for learning about functions because graphical constraining allows for learners to focus attention and cognitive energy on understanding functions. A combination of both strategies may allow learners to concentrate on functions, at first while self-explaining and then again while drawing, and/or allow learners to concentrate on one set of functions that are triggered from self-explaining and another set triggered from drawing. In either case, both study strategies are favorable for learning about functions.

When it came to learning about structures and pathways, the combination of the effects of self-explaining and of drawing appear much more complicated. Participants in the Combined group may have learned more than the SE group about structures and pathways because they were also able to draw. However, those in the Combined group also may have performed slightly worse than those in the Draw group because they had to first self-explain. What I am proposing by this is that the process of self-explaining first diverts attention from structures and pathways by encouraging explanations about functions. Subsequently, after self-explaining is complete, drawing brings attention to structures and pathways, but not to the same extent as drawing alone

would have. In other words, self-explaining before drawing simultaneously encourages learning about functions and discourages learning about structures and pathways.

Finally, these explanations remain consistent when considering the results of the blood flow task. The Draw group may have performed poorly on this task because mechanical automation from study session drawings compel students to overlay their image of the system onto the figure. It may be that the additional self-explaining that preceded each drawing during studying in the Combined group made students more inclined to consider how blood would flow through the system, discouraging that automation from forming for participants in that group.



## 5 PART 2: USE OF STUDY STRATEGIES

---

While Part 1 of this dissertation focused on comparing snapshots of student performance on formal assessments at different points in the study, Part 2 begins an exploration into the processes of using study strategies during independent study sessions that may have affected that performance. Analyses of video, audio, and drawing data collected during participants' independent study sessions are presented in two sections. First, self-explanations are considered from participants in the SE and Combined groups to examine relationships between the type and number of self-explanations that learners generated and their performance on the formal assessments. Second, drawings are examined from participants in the Draw and Combined groups to examine relationships between the type and number of drawings that learners generated and their performance on the formal assessments.

### 5.1 Analyses of Self-Explanations

---

The self-explanations generated by participants in the SE and Combined groups while studying passages 2 and 3 were analyzed. These passages were chosen for analysis for several reasons. First, they were the first passages that the participants studied without a researcher beside them. By the time a participant reached passage 2, he or she had already completed training (with practice), and self-explained the introductory passage of the CVS text with a researcher present for guidance or to answer any questions. At this point, participants should have been relatively comfortable with self-explaining and with what they were being asked to do with the remainder of the CVS text booklet. Passage 3, specifically, is an important passage in the booklet because it contains much of the information students need to understand the pathways of the two subsystems (i.e., the pulmonary system and the systemic system).

#### 5.1.1 Grain size

---

Self-explanations were coded from transcriptions of audio recordings from the SE and Combined groups (referring back to video data when necessary). These data were first divided into utterances, which were established by the structure of the explanation – that is, the form or purpose of what the learner said. Utterances were usually the length of a phrase and, at most, a sentence. This grain-size was chosen for several reasons. Chi, Leeuw, Chiu, & LaVancher (1994) analyzed their self-explanation data at the phrase level and additionally at a finer-grained level, finding the same pattern in both results. As explained in Chi (2000),

because we are mainly interested in knowledge inferences, coding at the grain size of the phrase seems to be more at the knowledge level, and the inference is more sensible, whereas coding SEs at a more fine-grained proposition level sometimes gets redundant. (p. 168)

This current study and analyses are also focused at the knowledge level. An analysis at a more fine-grained level would make it difficult to identify full explanations.

A larger grain-size was also considered, given that Chi, Bassok, Lewis, Reimann, & Glaser (1989) analyzed self-explanations at the level of a sentence or multiple sentences that referred to the same idea. In support of a larger grain-size, some phrases that might be considered a paraphrase at a more fine-grained level, could be coded to serve a larger purpose at a larger grain-size in the context of the whole explanation (e.g., as a contrast to another idea). However, within the SPF framework, the phrase-to-single-sentence level was most appropriate because one phrase about a structure, another about a pathway, and another about a function could all refer to

the same “idea” but, for our purposes, could be identified separately utterances. Thus, an utterance was identified in this study as a complete mention of a structure, pathway, and/or function that was as short as a phrase and as long as a sentence.

### 5.1.2 Defining self-explanations

---

Once transcriptions were separated into utterances, those that were self-explanations needed to be distinguished from those that were not. In line with coding schemes used in prior research on self-explanations (e.g., Ainsworth & Loizou, 2003; Chi et al., 1994; Chi et al., 1989), a self-explanation was defined in this study as an utterance that went beyond the information presented in the text, specifically, an inference. Take, for example, the following utterance generated after reading this line in the CVS text,

From passage 2: The septum divides the heart lengthwise into two sides.

SE #1: “Hm, I’m guessing it’s sort of like a tissue barrier.”

(#9119)

The sentence from the text does not explain what the septum is made of nor does it explain that it is solid. In SE #1, this student inferred that the structure of the septum is made of tissue and its function is implied to prevent blood from passing through. This knowledge is crucial to understanding the CVS because the septum prevents oxygenated blood in the left side of the heart from mixing with deoxygenated blood in the right side of the heart. As shown in this example, utterances were coded as self-explanations if they contained an inference.

In order to better understand which utterances qualified as self-explanations by this coding scheme, it is useful to understand what utterances *did not* qualify as self-explanations. These include monitoring statements, paraphrases, and translations. In other words, statements about understanding or not understanding content from the CVS text, such as “I don’t get it,” were not counted as self-explanations. Furthermore, paraphrases and translations, where learners rearranged words or substituted certain words with synonyms or sometimes their definitions, were also excluded as self-explanations. For example, the following utterance was not coded as a self-explanation:

From passage 3: In the lungs, carbon dioxide leaves the circulating blood and oxygen enters it.

SE #2: “in the lungs, when it arrives in the lungs, it gets gas exchange.”

(#4017)

In SE #2, the learner substituted the phrases about the movement of carbon dioxide and oxygen with the term “gas exchange.” This utterance does not add any new knowledge to what was described in the text, and would qualify as a paraphrase and not a self-explanations. Again, self-explanations were inferences that introduced new knowledge not presented in the text.

Once an utterance was identified as a self-explanation, it was further categorized by format. Self-explanations were initially coded by format according to the scheme created by Ainsworth & Loizou (2003). The only exception was the elimination of “principle-based explanations” from the scheme. As Ainsworth and Louizou explain, this format was scored if participants made reference to the principle of diffusion. However, given that, only self-explanations from Passages 2 and 3 were analyzed in this study and the passage explicitly about diffusion appeared much later in the booklet, our data rarely contained mentions of diffusion. The remaining self-explanation formats coded for were:

- Goal-driven: a purpose or aim was inferred of a particular structure or action
- Elaborative: information was inferred to add details to or expand on the topic. These included metaphors and analogies.
- Noticing coherence: an utterance was made about noticing an association between the current material and a previous idea (from the text or from prior knowledge).

The most common self-explanation format that participants generated were goal-driven. Coding for this format was often signaled by phrases that explained why an element in the CVS is structured a certain way, why blood flows to a particular place, or why a behavior is necessary. Note that these self-explanation formats were not mutually exclusive. Explanations were often about prior knowledge (noticing coherence) of a goal or an analogy (elaboration) that described a goal.

In addition to identifying formats among self-explanations, they were also coded for whether they mentioned structures, pathways, or functions. It is important to state that when a learner mentioned one of these categories in their self-explanation, it was not always an inference made about the structure, pathway, or function itself but instead served only as a reference or context for the self-explanation. Certainly, in regards to pathways, it would be difficult for a learner to make an inference about where blood flows when the text explicitly states the direction of blood flow. Hence, this coding was only of how often SPF categories were mentioned in self-explanations, and not how often they were the focus of self-explanations.

Mentioning functions happened to overlap with codes for goal-driven self-explanations. In fact, in the context of these analyses, a goal-driven explanation was synonymous with a description of a function. As such, all goal-driven self-explanations were coded as mentions of a function, and sometimes also as a mention of a structure or pathway. For example, a goal-driven self-explanation about why an element is structured a certain way was a self-explanation mentioning structure *and* function. Thus, all self-explanations that were goal-driven qualified as being about functions, but not all self-explanations about functions were goal-driven (e.g., those about behaviors).

### 5.1.3 Examples of self-explanations

---

For a better understanding of these self-explanation coding categories, let us consider some examples. Shown here is a sentence from the CVS text, followed by a response from one participant. In parentheses are the codes that were assigned to each utterance.

From passage 2: The right side pumps blood to the lungs, and the left side pumps blood to other part of the body.

*Participant #9749*

SE #3: “yeah the arteries on the side that bring the blood to the left side of the- to the heart are probably thicker” (Elaborative, structure)

SE #4: “because it has more pressure” (Goal-driven, function)

SE #3 was coded as an elaborative self-explanation and a mention of structure because it adds details about the structure of the left side of the heart, suggesting (beyond the information described in the text) that it is thicker. Then, in SE#4, the student suggests that the purpose of this thicker structure is to manage more pressure. This suggestion of a purpose is why this self-explanation was coded as goal-driven and a mention of function. Notice that the origin of the blood brought to the left side of the heart is not mentioned in this self-explanation, leaving the

path of blood flow unclear. Thus, the inferences generated in SEs #3 and 4 were only coded as mentions of the structure of the heart and the function that this structure serves.

In response to the same line in the CVS text, the following participant self-explained:

*Participant #4833*

Utterance: “The right side will pump blood to the lungs” (Pathway)

SE #5: “probably to get oxygen” (Goal-driven, function)

The first utterance did not qualify as a self-explanation because the information was explicitly stated in the text. However, given that this utterance set the context of the self-explanation that followed, it was coded as a mention of pathways. SE #5 was coded as a goal-driven self-explanation because it describes that the purpose of blood being pumped to the lungs is to acquire oxygen. This purpose was also coded as a mention of functions. Goal-driven self-explanations were the most common format generated by participants in this study.

Elaborative self-explanations were the second most common format found in the data. These often appeared as analogies to or metaphors for structures, pathways and functions, such as the following examples:

From passage 3: Two semilunar (s-l) valves separate the ventricles from the large vessels through which blood flows out of the heart.

SE #6: “And the semilunar valves probably look like half-moons”  
(Elaborative, structure; #4833)

From passage 2: In each side of the heart blood flows from the atrium to the ventricle.

SE #7: “And so it goes from down the top kind of like waterfalls, down a waterfall.”  
(Elaborative, pathway; #6148)

From passage 3: One-way valves separate these chambers and prevent blood from moving in the wrong direction.

SE #8: “One-way valves are something like sphincters because they only allow flow in one direction”  
(Elaboration, function; #4833)

These examples demonstrate how some learners were able to make associations between how some parts of the CVS can resemble other structures, pathways, or functions from everyday life. At first, “half-moons” in SE #6 may seem like a translation of “semilunar” but this student is making a connection between the parts of this word and the physical structure of the valve. This utterance was more than a simple substitution of synonymous words and is an inference about the name that can help the student understand the structure of these particular valves.

Elaborative self-explanations mentioning pathways were not generated as often as those mentioning structures and functions. Understandably, it is difficult to elaborate on repeated descriptions of blood moving from one place to another. SE #7 was one of only a few elaborative self-explanations about a pathway directly, explaining that blood flows from top to bottom through chambers in the heart just as water flows from top to bottom down a waterfall. Most other elaborative self-explanations that mentioned pathways did so simply to provide context.

Elaborative self-explanations mentioning functions were often elaborations about the functions themselves. In SE #8, for example, one-way valves were compared to sphincters as they each control movement through themselves. In general, an elaboration that learners

frequently made when self-explaining was about how blood vessels are similar to streets because they both control the direction of flow. Elaborations about functions often compared how something from everyday life behaved or had a purpose similar to something in the CVS.

The least common format of self-explanation that was identified in the data was noticing coherence. These may have been less common because these codes were assigned only to utterances that made the “noticing” part of the coherence explicit. In other words, participants were likely to recognize a lot of the material from prior knowledge or from previous passages, but did not explicitly state this recognition aloud. The following is an example of a self-explanation statement coded as noticing coherence.

From passage 3: Blood returning to the heart, which has high concentration, or density, of carbon dioxide and a low concentration of oxygen, enters the right atrium.

SE #9: “So the heart has a high concentration of CO<sub>2</sub>, which in the previous passage, was waste.”

(Noticing coherence; #0432)

While part of the information that is being described in SE #9 was simply taken from the text of the previous passage, this utterance is considered an inference because the connection between the passages was not explicitly stated in the text. Since, understanding that CO<sub>2</sub> is a waste product is paramount to understanding the CVS, making this type of inference is important.

In summary, transcriptions were divided into utterances, and all utterances were coded for whether or not they were self-explanations. Those that qualified as self-explanations were coded for format, consisting of goal-driven, elaborative, and noticing coherence. Finally, self-explanations were also coded for mentions of structure, pathways, and functions.

## 5.2 Results of Self-Explanation Analyses

---

The number of self-explanations generated by participants in the SE group and the Combined group were similar to each other, in format and in the number of times SPF categories were mentioned. T-tests determined that the difference between conditions in the number of overall self-explanations and self-explanation formats were not statistically significant. The differences in the number of times structures, pathways, and functions were mentioned in self-explanations were also not statistically significant.

The total and mean number of self-explanations generated by participants in the SE group and the Combined group are presented in Table 13. On average, learners who used the self-explanation strategy to study Passages 2 and 3 of the CVS text generated about 12 utterances that qualified as self-explanations. Also, on average, more than three of these self-explanations were goal-driven and more than three were elaborative. On average, functions were mentioned in self-explanations significantly more often ( $M = 5.80$ ,  $SD = 4.73$ ) than structures ( $M = 1.91$ ,  $SD = 2.75$ ;  $t(45)=6.96$ ;  $p < 0.001$ ) or pathways ( $M = 1.83$ ,  $SD = 2.41$ ;  $t(45)=7.79$ ;  $p < 0.001$ ).

**Table 13:** Coding of self-explanations (with standard deviations) for participants in the SE group and the Combined group

Group(s)		Coding of Self-Explanations by Format and SPF							
		Utterances	SEs	Goal-Driven	Elaborative	Noticing Coherence	S	P	F
SE (n=23)	<b>Total</b>	423	287	80	63	24	45	42	132
	<b>Mean</b>	18.39	12.48	3.48	2.74	1.04	1.96	1.83	5.74
	<b>SD</b>	11.67	9.56	2.87	2.42	2.34	2.74	2.62	5.42
Combined (n=23)	<b>Total</b>	467	273	85	95	19	43	42	135
	<b>Mean</b>	20.30	11.87	3.70	4.13	0.83	1.87	1.83	5.87
	<b>SD</b>	10.39	7.98	2.87	4.53	1.34	2.82	2.23	4.04
SE & Combined (n=46)	<b>Total</b>	890	560	165	158	43	88	84	267
	<b>Mean</b>	19.35	12.17	3.59	3.43	0.93	1.91	1.83	5.80
	<b>SD</b>	10.97	8.71	2.84	3.66	1.89	2.75	2.41	4.73

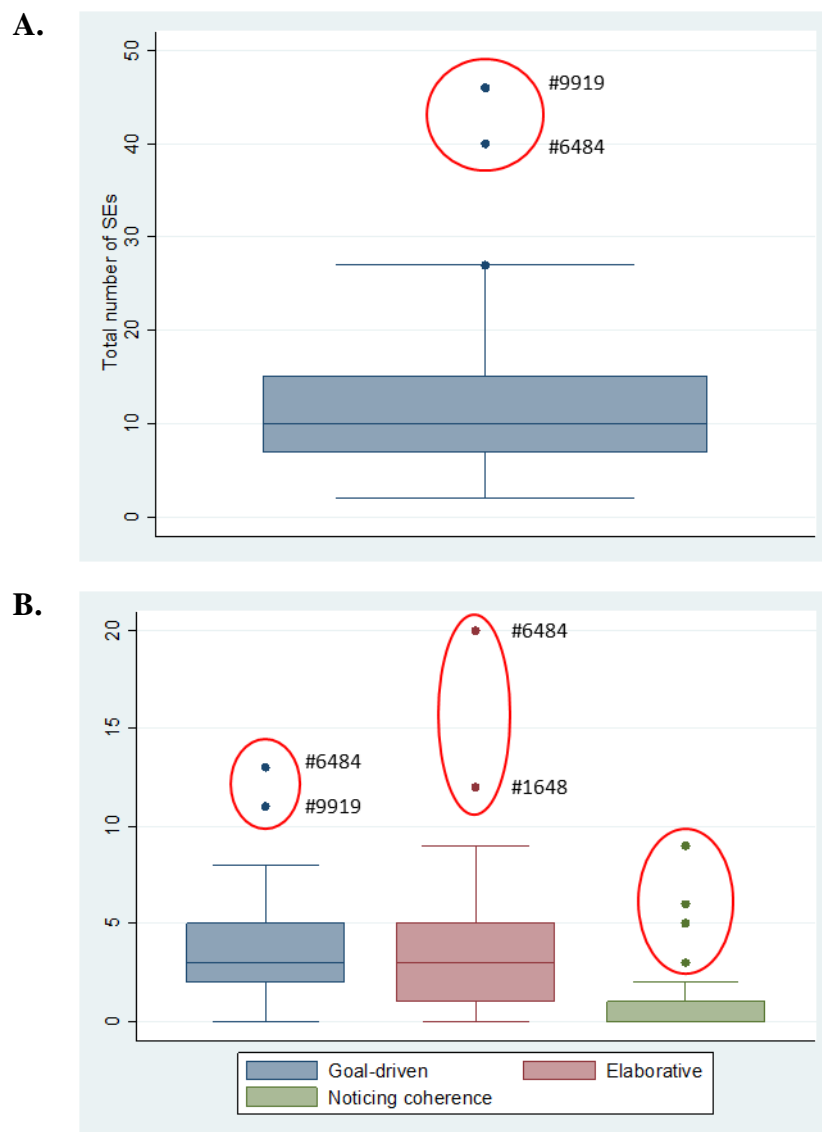
The following analyses and results are presented in two sections. First, self-explanation data from the SE group and the Combined group were pooled to examine how self-explaining, in general, might influence learning. Second, self-explanation data were analyzed separately for each group to examine how self-explaining may have been used differently by or had different effects on participants who only self-explained and those who self-explained and then drew.

### 5.2.1 Pooled: Self-explanations from the SE and Combined groups

Outliers from the combined data were identified using box and whisker plots shown in Figure 20. Data points circled in red were not included in any of the following analyses in this section. Two participants, #9919 from the SE group and #6484 from the Combined group, generated considerably more total self-explanations and goal-driven self-explanations than all other participants in the study and these data points were determined to be outliers in both categories. Participant #6484 also generated many more elaborative self-explanations than the other participants, along with participant #1648 from the Combined group. Very few participants generated more than two self-explanations coded as noticing coherence. Thus, the few participants who *did* were considered outliers. The noticing coherence category was not included in the following correlation analyses due to this lack of variation.

It is important to note that the data points that were dropped as outliers in these analyses were from participants who scored relatively well on the formal assessments. An analysis was also conducted with all self-explanation data points, including outliers, and are presented in [Appendix K](#). Many stronger correlations were found with the full data set, and the full data set may be representative of the true effects of self-explanation. However, the gap in data points between those that represent the majority of the participants and those that represent the few that self-explained a much greater amount calls for a more conservative analysis of only the majority, as they are a more representative sample. Therefore, the self-explanation data from these high-

self-explaining, high-performing participants are not included in the following correlation analyses.



**Figure 20:** Box and whiskers plots for **A.** total number of self-explanations and **B.** self-explanation formats (goal-driven, elaborative, and noticing coherence). Outliers circled in red were not included in the correlation analyses.

In order to conduct correlational analyses, formal assessment scores were collapsed into single scores regarding structures, pathways, or functions. Sub-scores from each SPF category were averaged from both QSA and QSB, on both the post and delayed assessments, and presented in the following tables as Total S, Total P, and Total F scores. These scores were then averaged to establish an overall Total SPF score. Drawing mental models scores were also averaged from short-term and long-term DMM gain scores. The correlation matrix for these formal assessment scores and total self-explanations and formats is presented in Table 14.

**Table 14:** Correlation matrix of formal assessment sub-scores and DMM scores versus total self-explanations and self-explanation formats for the pooled data from the SE and Combined groups.

Scores	Total No. of SEs	Self-explanation Formats	
		Goal-Driven	Elaborative
<b>Total SPF</b>	<b>0.30*</b>	<b>0.43**</b>	0.14
<b>Total S</b>	0.15	<b>0.34*</b>	0.06
<b>Total P</b>	0.25	0.27	0.12
<b>Total F</b>	<b>0.37**</b>	<b>0.51***</b>	0.12
<b>DMM</b>	0.10	0.12	0.22

\* $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\* $p < 0.001$

There were positive correlations across all categories, indicating that generating more self-explanations generally lead to higher formal assessment scores. There was a statistically significant moderate correlation between the total number of self-explanations generated by a participant and their total SPF score ( $r = 0.30$ ,  $n = 44$ ,  $p < 0.01$ ), mostly attributed to the strong correlation that exists between the goal-driven format of self-explanations and total SPF scores ( $r = 0.43$ ,  $n = 44$ ,  $p < 0.01$ ).

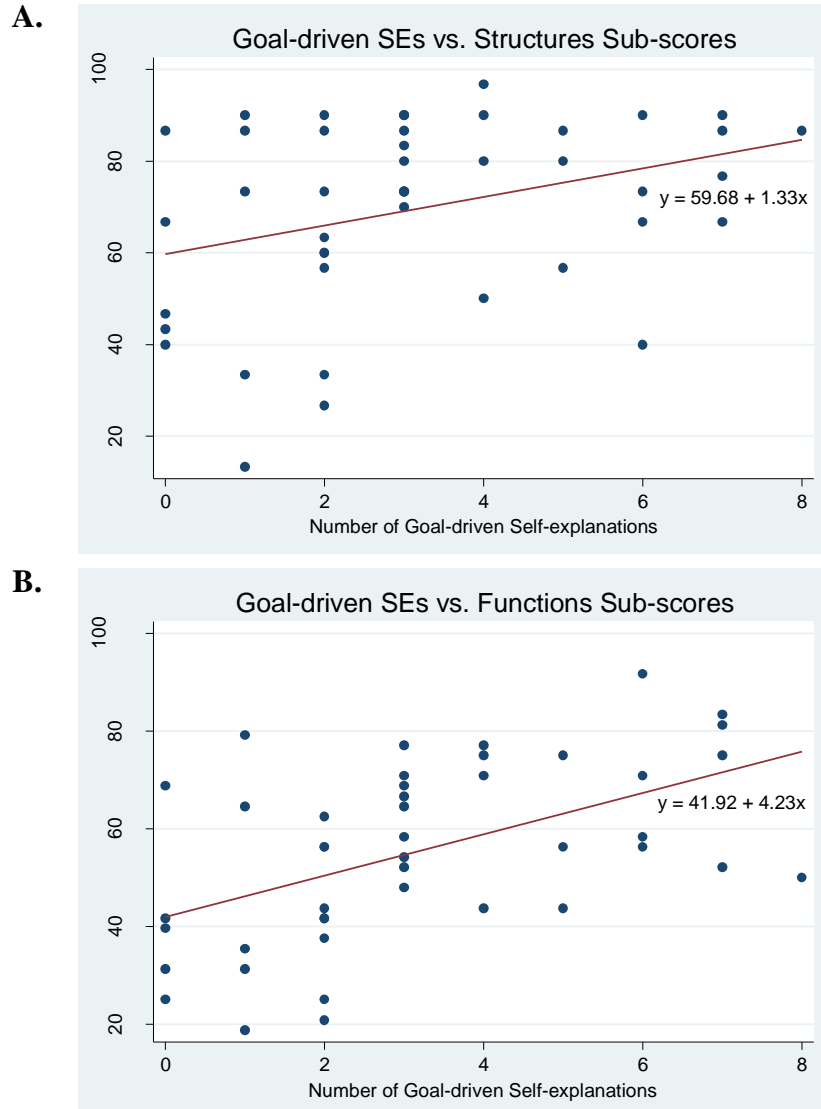
By SPF knowledge categories, the total number of self-explanations are correlated most strongly with functions sub-scores ( $r = 0.37$ ,  $n = 44$ ,  $p < 0.01$ ), likely attributed to the strong correlation between the goal-driven self-explanation format and functions sub-scores ( $r = 0.51$ ,  $n = 44$ ,  $p < 0.001$ ). Goal-driven self-explanations were also moderately correlated with structures sub-scores ( $r = 0.34$ ,  $n = 44$ ,  $p < 0.05$ ). Elaborative self-explanations were not significantly correlated with formal assessment sub-scores. Additionally, DMM scores were not significantly correlated to codes from any self-explanation format.

Linear regressions established that the number of goal-directed self-explanations was statistically significant at predicting structures sub-scores and functions sub-scores (structures:  $F(1,42) = 5.52$ ,  $p < 0.05$ ; functions:  $F(1,42) = 15.09$ ,  $p < 0.001$ ). Goal-directed self-explanations accounted for 9.52% of the variability in structures sub-scores and for 24.68% of the variability in functions sub-scores. The regression equations calculated were:

- *Predicted S sub-score (%) = 59.68 + 1.33 \* (# of goal-directed SEs)*
- *Predicted F sub-score (%) = 41.92 + 4.23 \* (# of goal-directed SEs)*

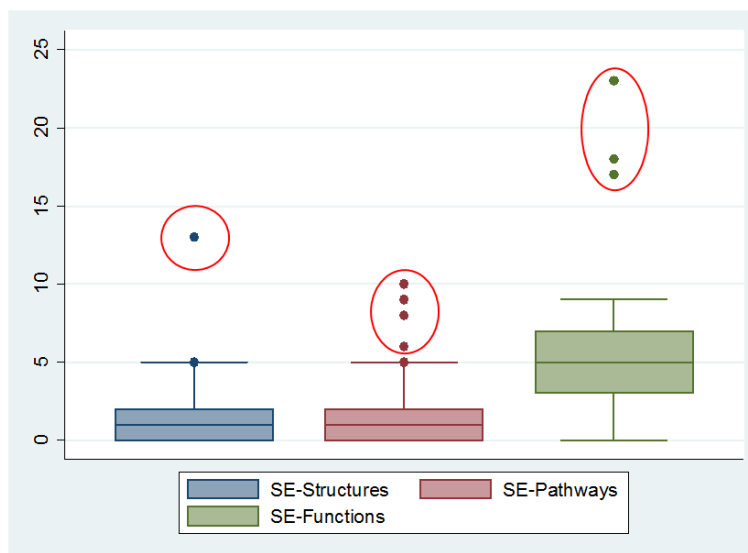
In other words, for every additional goal-directed self-explanation generated, a participant's formal assessment sub-scores are predicted to increase by 1.33 percentage points for structures (Figure 21A), and 4.23 percentage points for functions (Figure 21B).





**Figure 21:** Scatterplots with regression lines for **A.** structures sub-scores and **B.** functions sub-scores, predicted by the number of goal-driven self-explanations generated.

The codes for the mention of SPF categories in self-explanations were also used in correlation analyses with SPF sub-scores. Again, outliers were identified using box and whisker plots, and data points circled in red were dropped from the analyses (Figure 22). (An analysis conducted with all data points of mentions of SPF categories, including outliers, are presented in [Appendix K.](#))



**Figure 22:** Box and whiskers plots for mentions of structures, pathways, and functions in self-explanations. Outliers circled in red were dropped from correlation analyses.

Results of the correlation analyses conducted with the number of times SPF categories were mentioned in self-explanations and formal assessment scores are presented in Table 15. The matrix shows that total SPF scores were correlated with how often structures ( $r = 0.30$ ,  $n = 44$ ,  $p < 0.05$ ) and functions ( $r = 0.44$ ,  $n = 42$ ,  $p < 0.01$ ) were mentioned in self-explanations. Mentioning functions in self-explanations was significantly correlated with all three SPF sub-scores on (Total S:  $r = 0.34$ ,  $n = 44$ ,  $p < 0.05$ ; Total P:  $r = 0.30$ ,  $n = 42$ ,  $p < 0.05$ ; Total F:  $r = 0.50$ ,  $n = 42$ ,  $p < 0.001$ ). The correlation with Total F sub-scores was the strongest of the three. Also, mention of structures in self-explanations was correlated with pathways sub-scores ( $r = 0.35$ ,  $n = 44$ ,  $p < 0.05$ ), and with structures and functions sub-scores but were not statistically significant.

**Table 15:** Correlation matrix of formal assessment sub-scores and DMM scores versus mentions of SPF categories in self-explanations for the SE and Combined groups.

Scores	Mention of SPF categories in self-explanations		
	S	P	F
<b>Total SPF</b>	<b>0.30*</b>	0.14	<b>0.44**</b>
<b>Total S</b>	0.23	-0.02	<b>0.34*</b>
<b>Total P</b>	<b>0.35*</b>	0.12	<b>0.30*</b>
<b>Total F</b>	0.27	0.24	<b>0.50***</b>
<b>DMM</b>	0.13	0.04	0.18

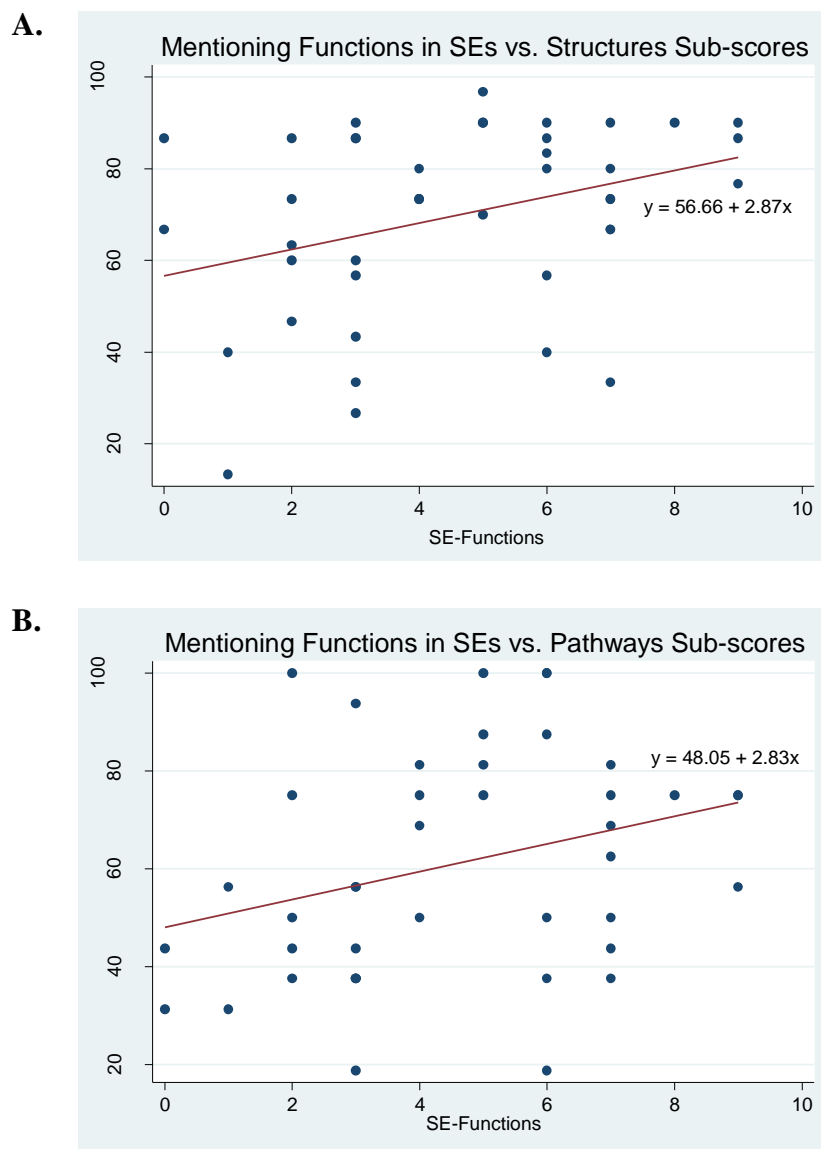
\* $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\* $p < 0.001$

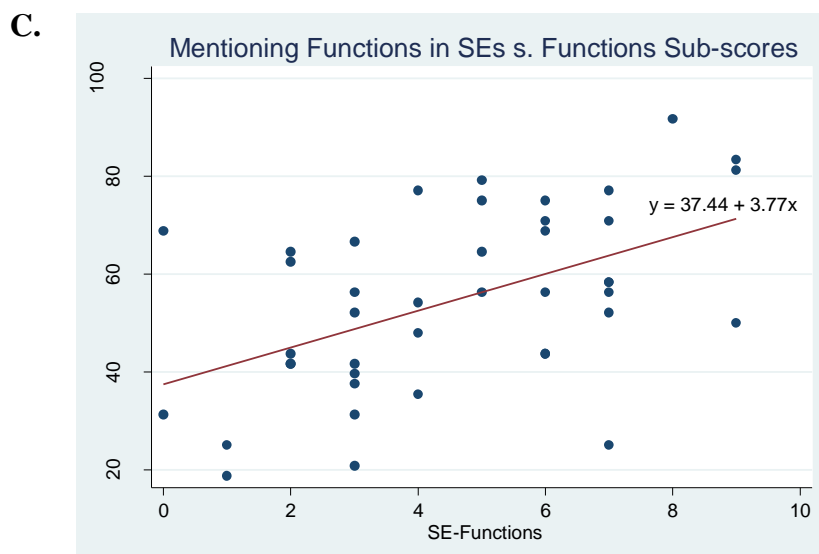
Linear regressions established that the number of times functions were mentioned in self-explanations was statistically significant at predicting SPF sub-scores (Total S:  $F(1,40) = 5.32$ ,  $p < 0.05$ ; Total P:  $F(1,40) = 4.01$ ,  $p < 0.05$ ; Total F:  $F(1,40) = 13.54$ ,  $p < 0.001$ ). Mentioning functions in self-explanations accounted for 9.53% of the variability in structures sub-scores,

6.85% of the variability in pathways sub-scores, and 23.42% of the variability in functions sub-scores. The regression equations calculated were:

- *Predicted S sub-score (%) = 56.66 + 2.87 \* (mentions of functions in SEs)*
- *Predicted P sub-score (%) = 48.05 + 2.83 \* (mentions of functions in SEs)*
- *Predicted F sub-score (%) = 37.44 + 3.77 \* (mentions of functions in SEs)*

In other words, for every mention of a function in a self-explanation, a participant's formal assessment sub-scores are predicted to increase by 2.87 percentage points for structures, 2.83 percentage points for pathways, and 3.77 percentage points for functions. Scatterplots and regressions lines for these data are presented in Figure 23, below.





**Figure 23:** Scatterplots with regression lines of **A.** structures sub-scores, **B.** pathways sub-scores, and **C.** functions sub-scores, predicted by the number of times functions are mentioned in self-explanations.

In summary, the total number of self-explanations a participant generated could be used to predict their functions sub-scores. This was largely due to the goal-driven self-explanations format. Goal-driven self-explanations were also statistically significant at predicting structures sub-scores, but to a lesser extent. Mentioning functions in a self-explanation was a strong predictor of functions sub-scores, and was significantly correlated with structures sub-scores and pathways sub-scores. Mentioning structures in self-explanations was also correlated with all three SPF sub-scores, but the only correlation that reached statistical significance was with pathways sub-scores.

### 5.2.2 Separately: Self-explanations from the SE and Combined groups

Correlation analyses were also conducted with self-explanation data from the SE and Combined groups separately. While identical training procedures were used to present self-explanation to participants in the SE group and Combined group, those in the Combined group were also trained and asked to draw after completing their self-explanations. If this added task affected the manner in which learners self-explained during their study sessions, these analyses may show different patterns of correlation between the two groups.

There was indeed a difference in the pattern of correlations of self-explanations and assessment scores between the SE group (Table 16A) and the Combined group (Table 16B). All correlations that were found to be statistically significant in each group were strong. In the SE group, the total SPF scores were positively correlated with the total number of SEs generated, and strongly with the goal-driven format, but these values did not reach statistical significance. The total number of self-explanations generated by a participant was correlated to pathways and functions sub-scores but were also not statistically significant. Goal-driven self-explanations were correlated to all three SPF sub-scores. Statistically significant and strong correlations were with pathways ( $r = 0.45$ ,  $n = 23$ ,  $p < 0.05$ ) and functions sub-scores ( $r = 0.54$ ,  $n = 23$ ,  $p < 0.01$ ). Elaborative self-explanations were weakly correlated with DMM scores.

With data from the Combined group, the total SPF scores were strongly correlated with the total number of SEs ( $r = 0.42$ ,  $n = 22$ ,  $p < 0.05$ ). The correlation with the goal-driven format

of self-explanations and total SPF scores was particularly strong ( $r = 0.49$ ,  $n = 22$ ,  $p < 0.05$ ). Correlations with the total number of SEs were moderate to strong with structures sub-scores and functions sub-scores. The correlation with functions sub-scores was strong and statistically significant ( $r = 0.46$ ,  $n = 22$ ,  $p < 0.05$ ). Goal-driven self-explanations from the Combined group were strongly correlated with structures sub-scores ( $r = 0.49$ ,  $n = 23$ ,  $p < 0.01$ ) and functions sub-scores ( $r = 0.47$ ,  $n = 23$ ,  $p < 0.05$ ).

**Table 16:** Correlation matrices of formal assessment sub-scores and DMM scores versus self-explanation formats from **A.** the SE group and **B.** the Combined group

<b>A. SE group</b>			
<b>Scores</b>	<b>Total No. of SEs</b>	<b>Formats</b>	
		<b>Goal-Driven</b>	<b>Elaborative</b>
<b>Total SPF</b>	0.23	0.40	0.00
<b>Total S</b>	0.02	0.33	-0.05
<b>Total P</b>	0.33	<b>0.45*</b>	0.03
<b>Total F</b>	0.31	<b>0.54**</b>	-0.04
<b>DMM</b>	-0.04	0.08	0.25

<b>B. Combined group</b>			
<b>Scores</b>	<b>Total No. of SEs</b>	<b>Formats</b>	
		<b>Goal-Driven</b>	<b>Elaborative</b>
<b>Total SPF</b>	<b>0.42*</b>	<b>0.49*</b>	0.28
<b>Total S</b>	0.37	<b>0.49**</b>	0.19
<b>Total P</b>	0.17	0.20	0.23
<b>Total F</b>	<b>0.46*</b>	<b>0.47*</b>	0.30
<b>DMM</b>	0.32	0.05	0.19

\* $p < 0.05$ ; \*\* $p < 0.01$

Overall, goal-driven self-explanations were strongly correlated with functions sub-scores in both groups. Goal-driven self-explanations were also more strongly correlated with pathways sub-scores than structures sub-scores for participants in the SE group, but the opposite was true for the Combined group. These patterns are reflected in the correlation values for total number of self-explanations. Lastly, elaborative self-explanations were not correlated with SPF sub-scores in the SE group, and were weakly correlated in the Combined group.

Correlation analyses conducted with the number of times SPF categories were mentioned in self-explanations also reveal different patterns between the SE group (Table 17A) and the Combined group (Table 17B). For the SE group, mentioning structures and functions, but not pathways, in self-explanations was associated with higher SPF sub-scores, most weakly with

structures sub-scores. The only correlation that was statistically significant was between mentioning functions in self-explanations and functions sub-scores ( $r = 0.45$ ,  $n = 21$ ,  $p < 0.05$ ).

For the Combined group, mentioning functions in self-explanations was strongly correlated with total SPF scores ( $r = 0.50$ ,  $n = 21$ ,  $p < 0.05$ ). While mentioning structures in self-explanations was generally weakly correlated with SPF sub-scores, mentioning functions was strongly correlated with structures sub-scores ( $r = 0.58$ ,  $n = 21$ ,  $p < 0.01$ ) and functions sub-scores ( $r = 0.58$ ,  $n = 21$ ,  $p < 0.01$ ). Mentioning functions was also moderately correlated with DMM scores for participants in the Combined group but not correlated at all for participants in the SE group. Finally, mentioning pathways in self-explanations was strongly correlated with functions sub-scores ( $r = 0.48$ ,  $n = 21$ ,  $p < 0.05$ ), but only in the Combined group.

**Table 17:** Correlation matrices of formal assessment sub-scores and DMM scores versus mentions of SPF categories in self-explanations from participants in **A.** the SE group and **B.** the Combined group.

<b>A. SE group</b>			
<b>Scores</b>	<b>Mention of SPF categories in self-explanations</b>		
	<b>S</b>	<b>P</b>	<b>F</b>
<b>Total SPF</b>	0.35	-0.02	0.38
<b>Total S</b>	0.23	-0.10	0.18
<b>Total P</b>	0.41	0.09	0.35
<b>Total F</b>	0.35	0.01	<b>0.45*</b>
<b>DMM</b>	0.19	0.09	0.00

<b>B. Combined group</b>			
<b>Scores</b>	<b>Mention of SPF categories in self-explanations</b>		
	<b>S</b>	<b>P</b>	<b>F</b>
<b>Total SPF</b>	0.28	0.31	<b>0.50*</b>
<b>Total S</b>	0.27	0.04	<b>0.58**</b>
<b>Total P</b>	0.33	0.14	0.19
<b>Total F</b>	0.20	<b>0.48*</b>	<b>0.58**</b>
<b>DMM</b>	0.11	-0.02	0.37

\* $p < 0.05$ ; \*\*  $p < 0.01$

In summary, different patterns of correlations were determined when self-explanation data from the SE group and the Combined group were analyzed separately. The total number of self-explanations, mostly attributed to the goal-driven format, was associated with higher functions sub-scores in both groups, higher pathways sub-scores in the SE group, and higher structures sub-scores in the Combined group. The same pattern existed with the number of times functions were mentioned in self-explanations, where this number was associated with functions sub-scores in both groups, pathways sub-scores in the SE group, and structures sub-scores in the

Combined group. The mention of structures was correlated with all three SPF sub-scores and most strongly with pathways sub-scores in both the SE and Combined groups, but none of these correlations were statistically significant. Mentioning pathways in self-explanations was not correlated with any of the SPF sub-scores in the SE group, but was strongly correlated with functions sub-scores in the Combined group. Finally, mentioning functions was not correlated to DMM scores in the SE group, but was in the Combined group.

### 5.3 Analyses of Booklet Drawings

---

Drawings that participants generated while studying particular passages of the CVS text were analyzed. Initially passages 2 and 3 were examined for consistency with the analyses conducted on the self-explanations for the same passages (presented in the previous section). However, because of the nature of passage 2, little variation existed in the coding of the drawings for this passage. Instead, another passage was chosen that was expected to show variation among participants' drawings.

The other passage chosen for analysis was passage 11, which described the details of the primary subsystems of circulation. This was the final passage in the booklet and was chosen because it presented a substantial amount of information that could be represented in a drawing in various ways. Passage 11 also contained information (especially about functions) that some students choose to include in their drawings while others did not. For example, from the sentence, The pulmonary artery is the only artery that carries deoxygenated blood, some participants indicated in their drawing that the pulmonary artery was distinct from all other arteries, while others only indicated that the pulmonary artery carried deoxygenated blood. These variations are in contrast to drawings made of passage 2, which is composed of six sentences, mostly about structures, for which all participants generated nearly identical drawings.<sup>5</sup>

A coding scheme was developed to analyze participant drawings of passages 3 and 11, based on the scheme used in Ainsworth, Galpin, & Musgrove (2007). This scheme accounted for all of the content presented in each line of the CVS passages and how it could be translated into a drawing. For example, from passage 11, the line Pulmonary Circulation is the movement of blood from the heart to the lungs and back to the heart, was interpreted to have two different items that could be included in a drawing: labeling "pulmonary circulation" and drawing an image of the loop of blood flow from the heart to the lungs and back.

The items evaluated in participant drawings for Passages 3 and 11 are listed in Figure 24, below. Each drawing was evaluated for the presence of each of these items in the form of a label and image. Given that different formats existed for the products of self-explaining, different formats for drawing were considered. From the coding scheme developed by Ainsworth, Galpin, & Musgrove (2007), it was apparent that labels were items often included in drawings, and could be considered a type of format of "drawing." Labels were considered written text, often just one or a few words in length, which identified an element or briefly explained a process. For example, rather than drawing the outline of two shapes to represent the two lungs, or a stick figure for a body, participant #9830 wrote "lungs" and "body" in her drawing (Figure 25). Labels

---

<sup>5</sup> The fact that passage 2 drawings did not vary much among participants, does not mean that learners are expected to have been processing the material in the same way or using the drawing study strategy to learn the information in the same way. This lack of variety in drawings only speaks to the limitations of this method of data collection. Participants were not asked to speak while they were drawing, and retrospective interviews were not conducted. Thus, the differences between the manners in which participants used the drawing strategy were not observable unless they appeared in the physical drawings that they generated.

were also sometimes counted as short textual descriptions of processes or functions, as listed in Figure 24.

### Passage 3

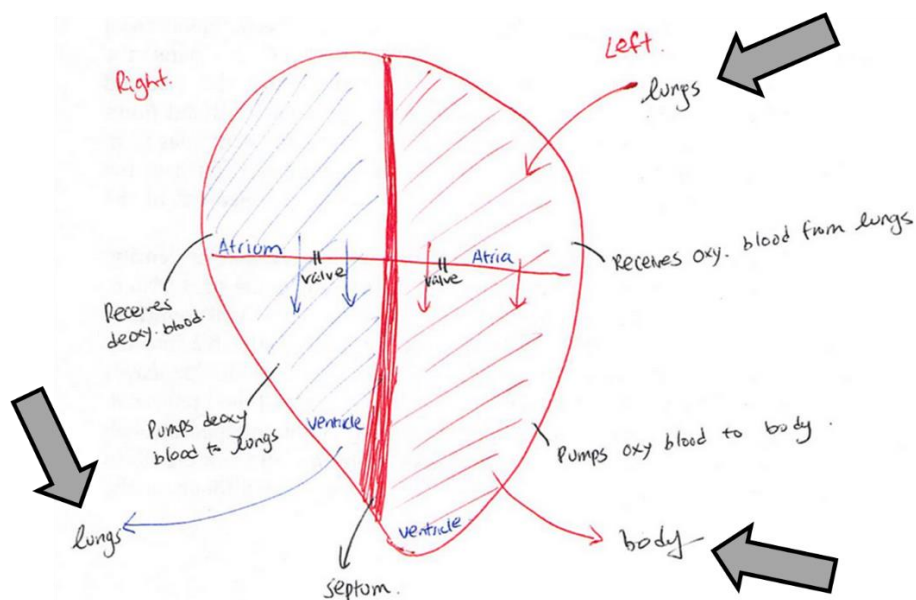
- Valves-prevent blood flow in wrong direction
- Valves separate chambers
- A-V Valve on right = tricuspid valve
- A-V Valve on left = bicuspid valve
- (I) Blood flows out of ventricles
- S-L Valve between ventricles and large vessels leaving heart
- (L) High concentration of CO<sub>2</sub>
- (L) Low concentration O<sub>2</sub>
- Deoxygenated blood returning to heart enters right atrium
- Pulmonary Artery
- Lungs
- In the lungs: CO<sub>2</sub> leaves
- In the lungs: O<sub>2</sub> enters
- (I) Oxygenated blood returns to left atrium of heart
- (I) Complete Loop
- (I) Blood flows through bicuspid/tricuspid valve into left/right ventricle
- Aorta
- Blood to rest of body from left side

### Passage 11

- (L) Pulmonary Circulation
- Movement of blood from the heart to the lungs and back to the heart
- Pulmonary Artery
- (L) the only artery that carries deoxygenated blood
- At the lungs the pulmonary artery divides into two smaller arteries one leading to each lung
- These arteries branch into arterioles and then into capillaries
- Oxygenated blood then flows into venules, which merge into the
- Pulmonary veins
- ...that lead to the left atrium of the heart.
- (L) Pulmonary veins are the only veins that carry oxygenated blood
- (L) Systemic Circulation
- Movement of blood to the rest of the body and then back to the heart.
- Blood from the left ventricle enters systemic circulation through the:
- Aorta.

**Figure 24:** Coding scheme items for drawings generated while studying Passages 3 and 11. Each drawing was evaluated for the presence of a label or image representing each of these items. Items marked with (I) were evaluated for the presence of an image only, and items marked with (L) were evaluated for the presence of a label only.





**Figure 25:** The three arrows point to examples of labels generated by participant #9830 while studying the CVS text.

If an item was not represented as a label in a drawing, it was represented as an image. An image was considered a geometric representation of one of the items on the list in Figure 24, generally in the form of a line, shape, or small collection of shapes. Note that the label and image categories were not mutually exclusive. There were many instances when an item described in the CVS text was represented as an image and also a label. For example, the septum in Figure 25 was drawn as a thick line dividing the heart vertically and was also labeled “septum.”

Some items in the coding scheme were evaluated only for their presence as images and not as labels. Blood flow and other types of movement were often only represented as images (i.e., arrows) and left unlabeled. For example, the item “Blood flows out of ventricles” was not labeled by any participant, and was either represented as an arrow pointed from the ventricle, away from the heart, or was not represented at all. On the other hand, some items in the coding scheme were evaluated only for their presence as labels and not as images. For example, the gases carbon dioxide and oxygen were always labeled as their names or as their abbreviations, “CO<sub>2</sub>” and “O<sub>2</sub>.” They were never drawn as images of gas molecules. Finally, in passage 11, the items “Pulmonary Circulation” and “Systemic Circulation” were only evaluated as labels because the details of the subsystems were considered evaluated in parts of other items.

#### 5.4 Results of Booklet Drawings Analyses

Participants in the Draw group generated more labels and images when drawing Passages 3 and 11 than participants in the Combined group (Table 18). On average, participants in the Draw group generated 16 labels and 21 images, while participant in the Combined group generated 13 labels and 18 images. These differences were statistically significant (Labels:  $t(45) = 3.22, p < 0.01$ ; Images:  $t(45) = 2.53, p < 0.01$ ).

**Table 18:** Coding of drawings (with standard deviations) for participants in the Draw group and the Combined group

Group(s)		Labels	Images
<b>Draw (n=23)</b>	<b>Total</b>	<b>372</b>	<b>474</b>
	<b>Mean</b>	16.15	20.61
	<b>SD</b>	3.57	2.58
<b>Combined (n=24)</b>	<b>Total</b>	304	422
	<b>Mean</b>	12.67	17.56
	<b>SD</b>	3.83	4.60
<b>Draw &amp; Combined (n=47)</b>	<b>Total</b>	676	896
	<b>Mean</b>	14.37	19.05
	<b>SD</b>	4.07	4.37

#### 5.4.1 Pooled: Drawings from the Draw and Combined groups

As with the self-explanation data, drawing data were first pooled from the Draw group and Combined group for correlation analyses and also analyzed separately by group. No outliers were identified by box and whiskers plots of labels and images of the pooled data (Figure 26).



**Figure 26:** Box and whiskers plots of labels and images from drawing data pooled from the Draw and the Combined groups

From the pooled data, labels were not correlated with SPF sub-scores and were weakly correlated with DMM scores. Images, on the other hand, were strongly correlated with structures sub-scores ( $r = 0.43$ ,  $n = 47$ ,  $p < 0.01$ ), and moderately correlated with functions sub-scores ( $r = 0.35$ ,  $n = 47$ ,  $p < 0.01$ ) and DMM scores ( $r = 0.38$ ,  $n = 43$ ,  $p < 0.01$ ).

**Table 19:** Correlation matrix of formal assessment sub-scores and DMM scores versus drawing labels and images for the pooled data from the Draw and Combined groups.

Scores	Drawings	
	Labels	Images
<b>n</b>	47	47
<b>Total SPF</b>	0.09	<b>0.39**</b>
<b>Total S</b>	0.11	<b>0.43**</b>
<b>Total P</b>	0.10	0.19
<b>Total F</b>	0.06	<b>0.35**</b>
<b>DMM</b>	0.26	<b>0.38**</b>

\* $p < 0.05$ ; \*\* $p < 0.01$

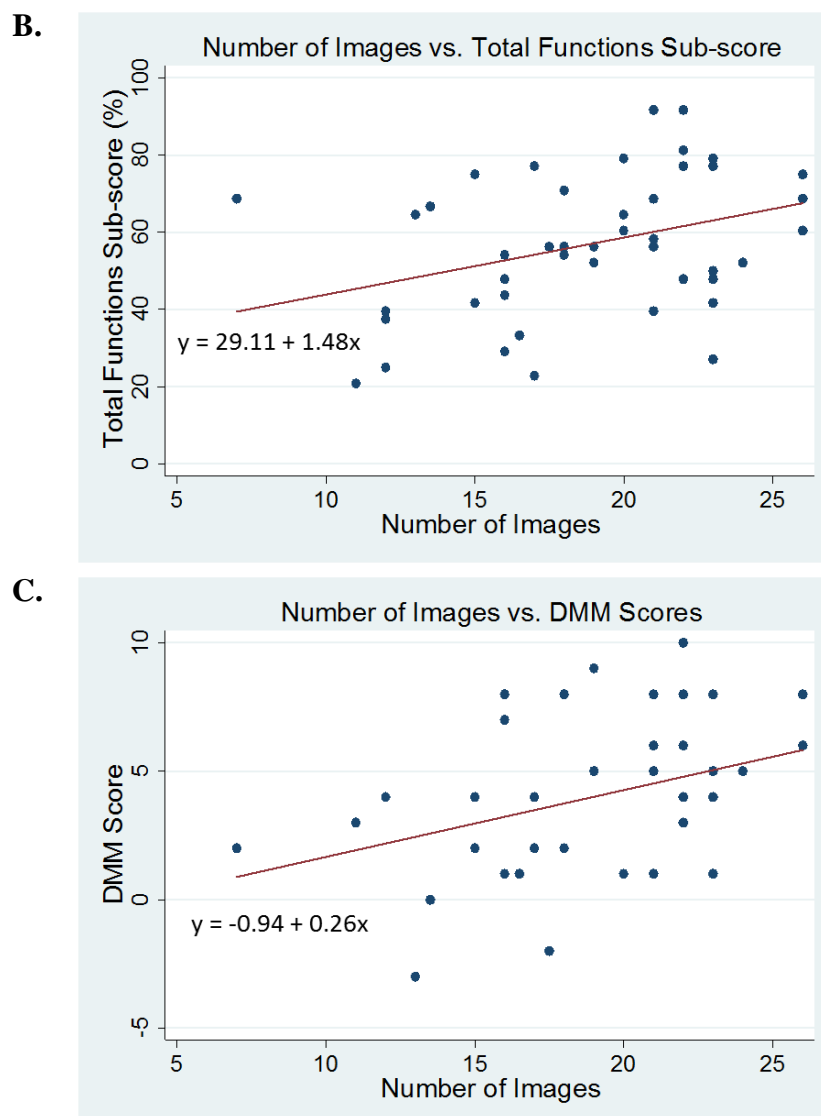
Linear regressions determined that the number of images that a participant generated was statistically significant at predicting structures and functions sub-scores, and DMM scores (structures:  $F(1,45) = 10.07$ ,  $p < 0.01$ ; functions:  $F(1,45) = 6.45$ ,  $p < 0.001$ ; DMM:  $F(1,41) = 6.90$ ,  $p < 0.001$ ). Images accounted for 16.47% of the variability in structures sub-scores, 10.59% of the variability in functions sub-scores, and 12.31% of the variability in DMM scores. The regression equations calculated were:

- *Predicted S sub-score (%) = 46.58 + 1.54 \* (# of images drawn)*
- *Predicted F sub-score (%) = 29.11 + 1.48 \* (# of images drawn)*
- *Predicted DMM score (units) = -0.94 + 0.26 \* (# of images drawn)*

In other words, for every additional image generated, a participant's formal assessment sub-scores are predicted to increase by 1.54 percentage points for structures, 1.48 percentage points for functions, and 0.26 points for DMM gains. The scatterplots and regression lines for these data are presented in Figure 27, below.

A.





**Figure 27:** Scatterplots and regression lines for **A.** structures sub-scores, **B.** functions sub-scores, and **C.** DMM scores predicted by the number of images generated.

In summary, the number of labels that participants in the Draw and Combined groups generated were not correlated to any of their formal assessment sub-scores. Images, meanwhile, could be used to predict structures sub-scores, functions sub-scores, and DMM scores.

#### 5.4.2 Separately: Drawings from the Draw and Combined groups

Correlation analyses were also conducted with drawing data from the Draw and Combined groups separately. Participants in each of these groups received identical training for drawing-to-learn, but those in the Combined group did not draw until they finished using the self-explanation strategy. This added task was likely to have an effect on the manner in which learners drew during their study sessions. If there was an effect, drawing data may reveal different patterns of correlations with assessment scores between the two groups, as it did with the self-explanation data.

Analyses of drawing data, conducted separately for the Draw group and the Combined group, revealed different patterns in correlations to formal assessment sub-scores (Table 20). In

general, all SPF sub-score correlations were stronger with drawing data from the Combined group than the Draw group.

When the data from the two groups were pooled, very weak or no correlations were found between the number of labels that learners generated and their formal assessment SPF sub-scores (Table 19). However, when we examine the data for labels from each group, separately, we see that this was a result of the negative correlations in the Draw group balancing the positive correlations in the Combined group (Table 20). A strong and statistically significant correlation was found in the Combined group between the number of labels generated and pathways sub-scores ( $r = 0.46$ ,  $n = 24$ ,  $p < 0.05$ ). The strongest correlation in the Draw group, on the other hand, was a moderate, negative one also between labels and pathways sub-scores (but did not reach statistical significance).

**Table 20:** Correlation matrices of formal assessment sub-scores and DMM scores versus the number of labels and images generated by participants in the **A.** Draw group, and **B.** Combined group.

**A. Draw group**

Scores	Drawings	
	Labels	Images
<b>n</b>	23	23
<b>Total SPF</b>	-0.22	<b>0.42*</b>
<b>Total S</b>	-0.07	<b>0.46*</b>
<b>Total P</b>	-0.31	0.16
<b>Total F</b>	-0.20	0.38
<b>DMM</b>	0.31	0.38

**B. Combined group**

Scores	Drawings	
	Labels	Images
<b>n</b>	24	24
<b>Total SPF</b>	0.39	<b>0.45*</b>
<b>Total S</b>	0.28	<b>0.48*</b>
<b>Total P</b>	<b>0.46*</b>	0.22
<b>Total F</b>	0.35	<b>0.43*</b>
<b>DMM</b>	0.19	0.41

\* $p < 0.05$

For correlations with the number of images that participants generated, the patterns between the groups were similar to each other and to those seen when the data were pooled from both groups. The strongest correlations were with structures sub-scores in each group (Draw:  $r = 0.46$ ,  $n = 23$ ,  $p < 0.05$ ; Combined:  $r = 0.48$ ,  $n = 24$ ,  $p < 0.05$ ), followed by functions sub-scores (Draw:  $r = 0.38$ ,  $n = 23$ ,  $p > 0.05$ ; Combined:  $r = 0.43$ ,  $n = 24$ ,  $p < 0.05$ ) and DMM scores. Finally, there were weak correlations with pathways sub-scores.

Overall, there were many strong and all positive correlations identified between formal assessment scores and the number of labels and images generated by participants in the Combined group. Some correlations were negative between scores and labels from the Draw group, but all were positive for correlations with images. In general, both labels and images were positively correlated to DMM scores in each group, with images having the stronger correlations.

## 5.5 Discussion of Self-explanation and Drawing Results

---

An examination of self-explanations generated by participants in the SE and Combined groups, and of drawings generated by participants in the Draw and Combined groups suggest several relationships between learning in certain SPF knowledge categories and self-explanation formats or drawing formats. Additionally, investigating the effects of each study strategy for each group, separately, further reveal relationships that differ between the groups.

The following discussion about the effects of self-explaining and drawing on learning must be qualified as correlational in nature. Linear regressions established what percentage of the variation in the data could be explained by these individual relationships but do not indicate causality, and leave the unexplained portion of the data unaccounted for. Many other unidentified factors are presumed to be involved in how these study strategies allow students to learn effectively, and may include intermediate factors responsible for the correlations seen in these results. However, these correlations are informative of how self-explaining and drawing may potentially be influencing the way students learn about the CVS from text and provide theories for further investigation in future research.

### 5.5.1 Self-explaining

---

The many positive correlations found between the self-explanation dataset and formal assessment scores suggest that self-explaining is an effective study strategy for learning about the CVS from text. These results support claims from prior research that the more self-explanations a learner generates, the more he or she tends to learn. Moreover, analyses conducted from the framework of SPF knowledge categories propose more specific effects of self-explaining on learning. Lastly, analyses conducted separately by study condition provide insight for explaining the differences seen in formal assessment results between the SE group and the Combined group.

#### *a. Self-explaining and learning, in general*

Self-explanation has been shown in prior research, in various ways, to be an effective strategy for learning. Chi et al. (1994) demonstrated that learners who generated a high number of self-explanations while studying the CVS text had greater pre- to posttest gain scores than learners who generated a low number. Ainsworth & Loizou (2003) established a strong positive correlation between students who generated more self-explanations from diagrams of the same CVS text and scores on post-tests. The correlation results from this study are consistent with these results, showing that the more self-explanations a learner generates while studying Passages 3 and 4 of the CVS text, the higher their SPF scores tend to be.

The Ainsworth & Loizou (2003) study also considered self-explanation formats in their analyses and found that the participants who self-explained from diagrams, tended to generate significantly more goal-driven self-explanations and score higher on post-tests, than participants

who self-explained from text. In another study, set in the context of learning about probability calculations, goal-driven self-explanations were one of the formats of self-explanation that had the strongest correlations with post-test scores, and elaborative self-explanations had a weak correlation (Renkl, 1997). The results from the current study also reflect these patterns.

Overall, this study was able to provide additional supporting evidence for the self-explanation effect: students who generate more self-explanations during learning tend to learn more than students who do not. It also shows that goal-driven self-explanations are a particularly useful format of self-explanation for learning.

*b. Self-explanation formats and SPF understanding*

As anticipated, generating a self-explanation of any kind was associated with higher scores in all three knowledge categories, but mostly functions. In particular, goal-driven self-explanations were the format that was most strongly associated with functions sub-scores. Given that goal-driven self-explanations are essentially about functions, and this format of self-explanation accounted for nearly 30% of learners' utterances, it makes sense that this relationship would exist.

Goal-driven self-explanations were also found to be significantly and positively correlated to structures sub-scores, and weakly correlated to pathways sub-scores. Some goal-driven explanations were specifically focused on the purpose of certain structures and pathways themselves, helping students to learn directly about components in these categories, but these types of self-explanations were not generated as often as the ones about functions. Perhaps it was necessary to understand structures and pathways to some extent because these knowledge categories were often used in the goal-driven self-explanations about functions. This would also help to improve structures and pathways sub-scores somewhat indirectly through self-explanation.

Mentioning SPF categories in self-explanations revealed the following relationships. Mentioning functions in self-explanations had similar positive associations as goal-driven self-explanations with all three SPF sub-scores. It is interesting to consider that mentioning functions in a self-explanation may improve a learner's understanding of structures and pathways. As proposed with goal-driven explanations, these associations may be a result of using the other SPF knowledge categories in the explanations about functions. The same reasoning would explain the fact that mentioning structures in a self-explanation is also associated with higher pathways and functions scores.

Mentioning pathways in self-explanations failed to show any association with SPF sub-scores when data from the SE and Combined groups were pooled. Remember that the pathways category was an addition to the coding scheme to account for responses to assessment questions that did not fall into a structure, behavior, or function category. An understanding of the path of blood flow is indeed important for learning about the CVS, but it is certainly specific to the CVS and cannot be generalized to all biological systems. Furthermore, it is unknown how important it is relative to the other knowledge categories.

It is difficult to generate a self-explanation specifically about pathways from a line in the CVS text that explicitly describes that pathway. Most self-explanations in this dataset that mentioned pathways were not explanations of the pathways themselves, but only mentions of

pathways as references or contexts for the rest of the explanation. The results suggest that simply mentioning pathways in a self-explanation is not enough to significantly improve learning about them, and perhaps pathways need to be the actual focus of the self-explanation in order to better promote learning. It may be that self-explanation is just not a good strategy for learning about the pathways of the CVS. Self-explanation appears to be the most useful for learning about functions.

The question remains: why are functions mentioned more often than structures and pathways when learners are self-explaining? Perhaps, to learners, the topic of functions appears more prominent in the CVS text, giving learners more material about functions to use in their explanations. Or maybe the opposite is true: functions are mentioned less explicitly, allowing learners more opportunities to ponder over the function of various elements in the CVS. This would be consistent with prior research that has demonstrated that learners tend to generate more self-explanations from reading text that is less coherent rather than more coherent (Ainsworth & Burcham, 2007). Then again, it may just be the nature of self-explanations themselves, being more conducive to topics about functions as learners often try to infer causality when they generate explanations (Keil, 2006).

When considering functions in the CVS, both *cause and effect* can easily relate to how and why anything happened in the system. Structures, on the other hand, can definitely have *effects* on the system, but what *causes* a component to be structured a certain way is extraneous information. Finally, blood flowing from one place to another (pathways) only has an *effect* when the destination is capillaries where nutrients and waste can be exchanged, and what *causes* blood to flow from one place to another is always because of the same reasons: valves and differences in pressure. Self-explanations about causality may be the easiest to generate, and those may most naturally be about functions.

Note that I am not proposing that learners attempt to generate as many self-explanations about structures and pathways as they do about functions in order to understand all SPF categories in the CVS better. Such an attempt may be more strenuous than effective. Rather, the patterns presented among these self-explanation characteristics and performance on formal assessments suggest that using the self-explanation strategy to learn about the CVS from text seems to place an inherent emphasis on functions. Considering that function is the knowledge category that is most often overlooked by learners (Chi et al., 1994; Hmelo-Silver & Pfeffer, 2004), it makes sense that self-explanation has been repeatedly shown in the literature to be an effective strategy for learning.

### *c. Self-explaining vs. Self-explaining and Drawing*

On average, participants from the SE and Combined groups spent approximately the same amount of time self-explaining during their study sessions, and generated approximately the same number of self-explanation formats. However, those in the SE group averaged lower SPF sub-scores than participants in the Combined group. While it is expected that this difference is mostly due to the extra drawing task that participants in the Combined group completed after self-explaining, some differences were found in correlations between self-explanation formats and formal assessment scores. Meaning, while participants in the Combined group did use an additional study strategy to self-explaining, correlations were identified between assessment



scores and things these participants did only during the self-explanation portion of their study sessions.

For participants in the SE group, goal-driven self-explanations were shown to be moderately to strongly correlated with all SPF sub-scores, and elaborative self-explanations showed no relationship with any SPF sub-scores. For participants in the Combined group, however, the relationships were different. Goal-driven self-explanations were most strongly correlated with functions and structures sub-scores, and very weakly with pathways sub-scores. A very similar pattern was seen with correlations between mentioning functions in self-explanations and SPF sub-scores in each group. In the SE group, mentioning functions was moderately-to strongly correlated with functions and pathways sub-scores and very weakly correlated to structures sub-scores. In the Combined group, mentioning functions was even more strongly correlated with functions and structures sub-scores and very weakly correlated to pathways sub-scores.

Overall, more strong and significant correlations were found in the Combined group between self-explanation characteristics and formal assessment scores compared to those found in the SE group. This difference suggests that the added drawing task assigned to participants in the Combined group may have influenced the way that they self-explained or the way that self-explanation influenced the way that they learned. If there were no differences in how self-explanations were associated with SPF sub-scores between the SE and Combined groups, it would be more likely that the differences in their average formal assessment scores were due only to the additional drawing task. However, the differences found in correlation patterns suggest that the differences in assessment scores were due (at least in part) to self-explaining.

The anticipation of drawing may affect the manner in which participants use the self-explanation strategy. Drawing was presumed to focus learner attention on structures and pathways more so than self-explaining because of the inherent requirement to include these features in one's drawings. The fact that certain types of self-explanations were more strongly associated with learning about structures in the Combined group than the SE group suggests that drawing may have indeed focused learner's attention on structures. However, instead of focusing attention on structures only while drawing, perhaps it was while self-explaining, in anticipation of drawing.

Student learning of pathways is more difficult to assess. The straightforward and simple nature of the text that describes information in this category make self-explanations about pathways difficult to generate. With structures, participants can often explain functions and structures simultaneously, describing why elements are structured in particular ways and how the structure helps it to fulfill its purpose. It is sensible that in anticipation of drawing, participants in the Combined group make note of structures (in addition to functions) while they are self-explaining. Learning about pathways, on the other hand, may be left to occur during the drawing task for participants in the Combined group because they are difficult to generate self-explanations about. Whereas students in the SE group may have tried to self-explained more generally about all categories.

In summary, self-explanation has been shown to be an effective strategy for learning about the CVS. Results from this study suggest that goal-driven self-explanations are the most effective format for learning. Self-explanations were also found to be most strongly associated

with learning about functions, and I presume this is because self-explanations are useful for generating inferences about how and why components work. Furthermore, self-explanations generated by participants in the Combined group, compared to those in the SE group, may be more focused on structures that they will need to draw for the subsequent study strategy they will use. However, it may be more difficult for them to focus self-explanations on the pathways that will need to be drawn because pathways presented in the text often do not require explanation.

### 5.5.2 Drawing

---

Results from the analyses of student drawings support prior research demonstrating that drawing is an effective study strategy for learning from text. The drawing study strategy was originally hypothesized to have a stronger effect on learning about structures and pathways than self-explaining, and is found to be partially true. Correlation analyses conducted with the number of labels and images participants generated in their drawings reveal potential differences between how the drawing strategy was used by participants in the Draw group and by participants in the Combined group.

#### *a. Drawing and learning, in general*

In many contexts, drawing has been shown to be an effective study strategy for learning (e.g., Ainsworth et al., 2007; Cox & Brna, 1995; Van Meter, 2001). Given that the content being learned and the extent to which a drawing activity is scaffolded greatly effects the benefit that drawing can have on learning (Van Meter & Garner, 2005), prior studies have not been able to establish a consistent measure for drawing characteristics that can effect learning. In this study, the number of labels and images that participants produced were identified to serve as drawing “formats” that might have different effects on learning, as different self-explanation formats were found to have. The coding scheme concentrated on content from the CVS text, quantifying the number of items from the text that could be represented and identifying them in student drawings. Coding for images that represent specific content may be the start to a reliable method for assessing student drawings when examining the effects of drawing on learning.

With data pooled from the Draw group and the Combined group, significant correlations were found between most formal assessment sub-scores and the number of images generated, but no correlations were found with the number of labels generated. Certainly, generating images is a more complex task for a learner than generating labels, since labels are often just words copied directly from the CVS text. Images cannot be copied from the text and must be somehow transformed from text form into image form.

One study that looked at representational skills compared the abilities of expert and novice chemists to make representational transformations. They found that “[e]xperts were significantly better than novices at transforming a given representation into a chemically meaningful representation in another form (Kozma, 2003; p.208). In this study, every image that a participant generated was essentially a transformation of information from a textual form into a visual one. Thus, learners who transformed information from the CVS text into images more often (i.e., generated more images in their drawings) may have learned more effectively.

#### *b. Drawing formats and SPF understanding*

The number of images a student generated was found to have the strongest correlation with structures sub-scores. Considering that drawing the CVS makes a learner’s understanding of

structures explicit, since it is necessary that they include these components in their drawings, it is not surprising that learners who used the drawing strategy had improved understanding of structures. Drawing seems to support metacognition, particularly about structures, by making obvious what they think certain structures look like and how they think they are arranged.

Drawing was also anticipated to be particularly useful for improving drawing mental models. In this study, both labels and images were positively correlated with DMM scores. Clearly, many passages in the CVS booklet contained parts or all of the information that participants would include in their drawings on the formal assessments. In essence, drawing during studying, to some extent, as practice for the DMM task on the assessments. During their study sessions, participants who drew could have tested different ways to transform the information in the text into images or different arrangements of the CVS components, perhaps allowing them to discover combinations that allowed them to include more information in their drawings. Practice labeling also seems likely to help with remembering names of components and clusters or chunks of them in their drawings. These reasons might explain why participants who drew during their study sessions were likely to see improvement in their DMM scores.

Learner understanding of functions was anticipated to improve if students viewed their drawings as models, as they were taught to during training. In this sense, their drawings could be used to simulate a working system and imagine how the various components function individually and together. Learners may have also been considering the functions of various components as they added each of them to their drawings.

The weakest correlation with the number of images drawn was with pathways sub-scores. Given that the Draw and Combined groups each had relatively high pathways sub-scores, the lack of a strong correlation with the number of images generated suggests that these groups learned about pathways in some other manner related to drawing. While images of some pathways were identified in drawings as lines or arrows between larger structures, a majority of drawings were occupied with details of these larger structures. It may be that the number of these structures is not associated with a learner's understanding of pathways because they emphasize the structures more than the paths between them. This would partially explain why such a weak correlation was found between the number of images generated and pathways sub-scores.

An understanding of the systemic and pulmonary sub-systems from a drawing only requires an image of the heart (divided into four chambers), the lungs, and the body. Even if images for components found between these major structures were not included in a drawing, the learner may still know of these structures along the path of blood flow and simply did not include these details in his or her drawing. Thus, it may be that additional images represented in a drawing beyond these three major structures does not significantly effect a learner's understanding of pathways. Moreover, it may just be the visual aid of the CVS, no matter how detailed, that helped participants in the Draw and Combined group learn about pathways.

### *c. Drawing vs. Self-explaining and Drawing*

The correlations found between assessment scores and the number of images drawn persisted when the data were analyzed separately by group. The pattern of correlations were also roughly the same between the two groups. The strongest correlation was between images and structures sub-scores, followed by correlations with functions and DMM sub-scores. Lastly, images were very weakly correlated with pathways sub-scores in each group. All correlations were slightly stronger in the Combined group than the Draw group. The similarity in these patterns for the two groups suggest that images have the same impact on learning for participants

in each group. While each image can be represented very differently among students, the act of generating an image, in general, may be enough to improve learning for all students. Furthermore, this may be related to the act of transforming the textual representation into a visual one, as mentioned earlier.

With labels, on the other hand, analyses conducted with drawing data from each group separately, revealed different patterns from the pooled data. Between labels and assessment scores, negative correlations were found with data from the Draw group, while positive correlations were found for the Combined group. It is unclear why labeling would have opposite effects in these two groups, but perhaps the act of labeling in the Combined group is made more meaningful after participants complete their self-explanations. Participants in the Draw group immediately generate their drawings upon reading the passage, sometimes simultaneously while they are reading. Thus, they likely spent less time processing information before generating labels and images for their drawings. On the other hand, participants in the Combined group self-explained the content before drawing and may have more knowledge associated with their labels. The same effect might not apply to images because the difficult process of transforming text into images may mute any residual effects of self-explaining before drawing.

### 5.5.3 Self-explaining vs. Drawing

---

Without a formal control group in this study, it is difficult to make claims about the extent to which participants in one group learned relative to a baseline. However, Part 2 of this dissertation allows us to make claims about participants relative to each other. We see that a participant who self-explains more, tends to learn more than one who self-explains less. Goal-driven self-explanations are a particularly useful format for learning about functions. Also, a participant who draws more, tends to learn more than one who draws less. Images are a particularly useful format of drawing for learning about structures. In addition to comparing participants within study strategy groups, we are also able to make some comparisons across them.

The SE group, on average, performed the worst of the groups on assessment items regarding structures. I believe self-explanation does indeed promote learning about structures to some extent but not as much as drawing or with drawing. This is supported by the fact that self-explanations *were* correlated with structures sub-scores, just weakly or moderately. Drawing (quantified by the number of images generated), on the other hand, was strongly correlated with structures sub-scores in both the Draw group and the Combined group. Together, these results suggest that drawing supports student learning of structures better than self-explanation does.

Functions sub-scores tended to be rather equal across the SE and Draw groups. This can be explained by the fact that the number of goal-driven self-explanations and images generated were both strongly correlated to functions sub-scores. Thus, self-explaining and drawing may be equally useful for learning about functions. Interestingly, participants in the Combined group appeared to average only slightly higher on their functions sub-scores than participants in the SE group and Combined group. It appears that learning about functions using both strategies may be rather redundant. However, the slight increase in score may be attributed to the advantage of solidifying understanding of a function the time the second strategy is used to learn it, or perhaps to different sets of functions that are elicited by using each strategy to learn.

Finally, the SE group tended to score lower than the Draw and Combined groups on pathways sub-scores. While self-explanations were weakly correlated with pathway scores, and images and labels were not consistently correlated to pathways sub-scores across groups, I believe that drawing is more effective than self-explaining to learn about pathways. The benefit may not appear in the form of a correlation because the number of images in a drawing is not directly related to understanding pathways, but a visual tool that a learner can use to simulate the blood flow is likely to help him or her understand how blood travels through the CVS. The data, or analyses conducted thus far, in this study just do not capture the aspect of the drawing strategy that accounts for this learning.

## 6 PART 3: CASE STUDIES

---

Thus far, we have considered much of the quantitative data collected in this study. We compared scores from the formal assessments and counted types of self-explanations and drawings that participants generated. In this final part of the study, we will take a qualitative look at the way two different participants used the self-explanation and drawing study strategies to learn about the CVS. One participant, who greatly improved his understanding from the pre-test to the post and delayed tests, showed evidence in his self-explanations and drawings of making connections in the form of integrating material to develop a coherent model. The other participant, who improved an average amount across tests, also showed evidence of making connections but of lower quality.

### 6.1 Analyses of Self-explanations and Drawings

---

In order to evaluate how self-explaining and drawing may influence learning, we need to look at the quality of the products of these study strategies. Part 2 of this dissertation considered various formats of self-explanations and drawings, but format is not informative of quality. That is to say, one goal-driven self-explanation or image can be more meaningful than another. Furthermore, while quality in everyday life is usually judged based on the manner in which the product is presented to an audience, here we must shift our standards a bit to account for the fact that participants in this study were asked to *self-explaining* and draw for themselves. Therefore, we must be sensitive to how much “interpretability” factors into our judgements about quality for learning. For this analysis, we will characterize the quality of these products in terms of how useful they could be for understanding the CVS, its many parts and how those parts work as a system.

#### 6.1.1 Quality of Self-explanations

---

The coding scheme used in Part 2 of this dissertation allowed us to separate self-explanations from non-self-explanations (e.g., paraphrases and monitoring statements). However, it did not allow for judgements to be made about the quality of the self-explanations, which certainly ranged widely in the verbal data. Analyses of the quality of self-explanations might address why students who generated more tended to perform higher on the post and delayed assessments. Was it simply due to the sheer number of goal-oriented self-explanations a learner generated or was it more specific about what information a learner used in their self-explanation?

As mentioned in the last chapter, some prior studies have examined types of self-explanations that students have generated, even specifically about the CVS. However, an analysis of the quality of these self-explanations, in terms of their usefulness for learning the content, has yet to be conducted. Thus, we borrow criteria used for examining quality, in these terms, from studies indirectly related to self-explanations.

Brewer, Chinn, & Samarapungavan (1998) describes a list of attributes that scientists tend to use to judge the quality of explanations. Given that our participants are not scientists and were not asked to generate explanations about empirical data, not all of the attributes on this list were relevant to this study. However, most were directly applicable to or easily modified to fit our self-explanation data. In general, a self-explanation of higher-quality can be characterized by the following:

- **Accuracy:** It is consistent with the information presented in the text.

- **Scope:** It expands the range of information presented in the text to include a larger perspective of the CVS.
- **Consistency:** The information presented in the self-explanation is internally aligned and aligned with other, recent self-explanations.
- **Simplicity:** It relies on few assumptions and hypotheses.
- **Precision:** It includes details specific to the topic.

Since each of Adam's self-explanations regularly met all of these criteria for high-quality explanations, it would be difficult to discuss his examples by criteria, separately. Instead, his self-explanations will be described by these attributes in general, and also by their effects. In other words, we will discuss what purpose each of Adam's self-explanations may have served in terms of him learning about the CVS from the text. The effects of Beth's self-explanations, on the other hand, were unclear. Thus, her examples will mostly be discussed in terms of this list of attributes.

### 6.1.2 Quality of Drawings

---

In the analysis of participant drawings in the last chapter, the coding scheme used only allowed us to count the number of times pieces of information from the text were represented in student drawings, in the form of labels and images. While this was informative of how much detail students included in their drawings, it was not necessarily informative of how useful the drawings were for the student to learn about the CVS. Thus, a closer examination of individual students' drawings aim to characterize their quality.

Note that since participants were asked to draw for themselves and not for the researcher or anyone they might wanted to imagine, not all drawings were expected to be fully interpretable by anyone other than the student who drew them. Therefore, in a drawing, some meaning and value may not be distinguishable to an observer but may be obvious to the student who drew it. Unfortunately, we do not have the data to identify such instances. Instead, we can only make minimal assumptions about the interpretability of a drawing that might be related to learning. For example, if one drawing is easier to interpret than another it may be an indication that the student who drew the former had a clearer or better formulated mental model of the system than the student who drew the latter.

Keeping in mind that students were drawing for themselves, we can consider characteristics, other than interpretability, that might describe higher-quality drawings. After examining the drawings generated by the two case study participants, attributes were identified to differentiate higher-quality drawings from lower-quality drawings. The identification of these attributes was strongly influenced by the lists described in diSessa, Hammer, Sherin, & Kolpakowski (1991) and diSessa (2002). In those studies, patterns were found in the types of judgements students made either while generating representations or while critiquing representations, ultimately amassing lists of criteria young students tend to use to make judgements about representations. For this study, however, we will judge the quality of the representations ourselves based on how useful they can be for learning about the CVS. These lists and the discussions of them in these papers were used as starting points for this analysis.

We will consider the following attributes of participant drawings that tend to describe higher-quality drawings:

- **Content focus:** Somewhat related to “completeness” (diSessa et al., 1991) and “epistemic fidelity” (diSessa, 2002), drawings tend to accurately represent all relevant information from the associated passage. The emphasis here is on the *relevant* information, these drawings tended to focus on the main ideas presented in the passage.
- **Scope:** Drawings tend to have a narrow physical scope of the CVS components described in the passage. Scope is very much associated with content focus, but scope applies more to the sense of whether drawing is zoomed-in or zoomed-out on the CVS, whereas content focus is about what information is the focal point of the drawing. That is to say, a drawing can have a wide scope and include the whole body but the content can be focused on a specific component in the center of the drawing.
- **Level of abstraction:** Drawings are more abstract, not always representing components in their correct anatomical shapes or arrangements. Drawings that were less abstract were more realistic, concrete, or figurative.
- **Representation of motion:** Drawings have a clear system for representing movement and cycles.

## 6.2 Identifying Participants for Case Studies

---

A closer examination of students’ self-explanations and drawings aim to illuminate how they used these study strategies to learn about the CVS, what aspects of these strategies may be more or less effective for learning, and what the products of these strategies can tell us about students’ understanding. Two individuals from the Combined group were selected for this analysis in order to examine and compare their uses of both study strategies. These two students were selected because they scored similarly on the pre-test and spent approximately the same amount of time with the CVS text during their independent study sessions (Figure 28). One student, who we will call Adam, performed relatively well on the post and delayed tests. The other student, who we will call Beth, performed near the average for all participants, and not as well as Adam (Figure 29). In studying these two particular students, we aim to identify differences in the way they used the self-explanation and drawing strategies to arrive at different learning outcomes from similar beginnings.



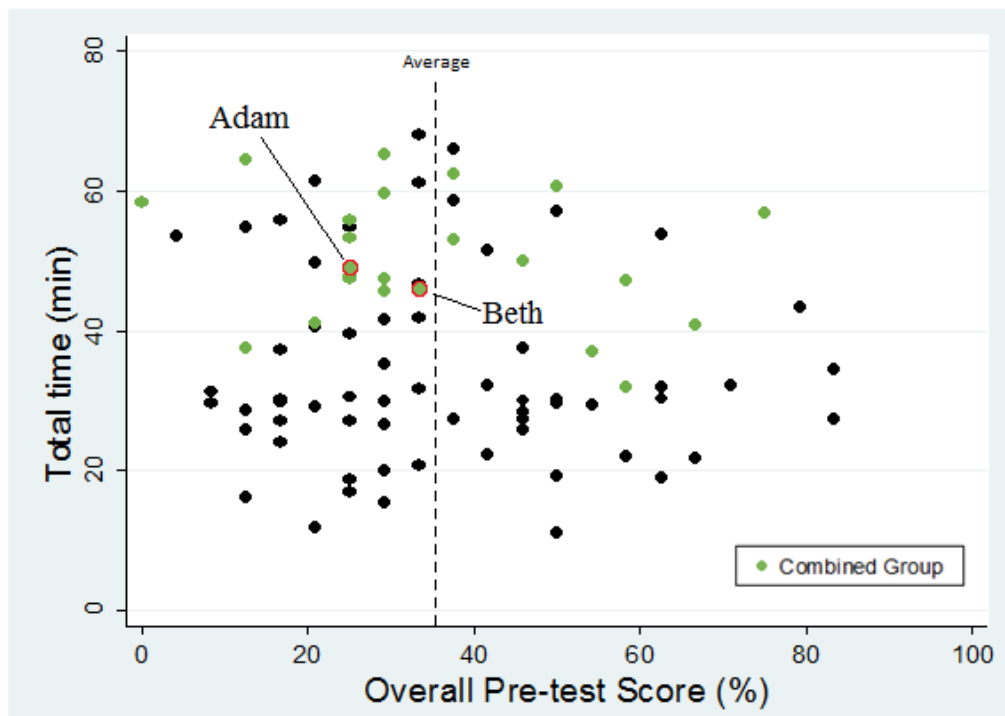


Figure 28: Adam and Beth had similar pre-test scores and total times spent studying.

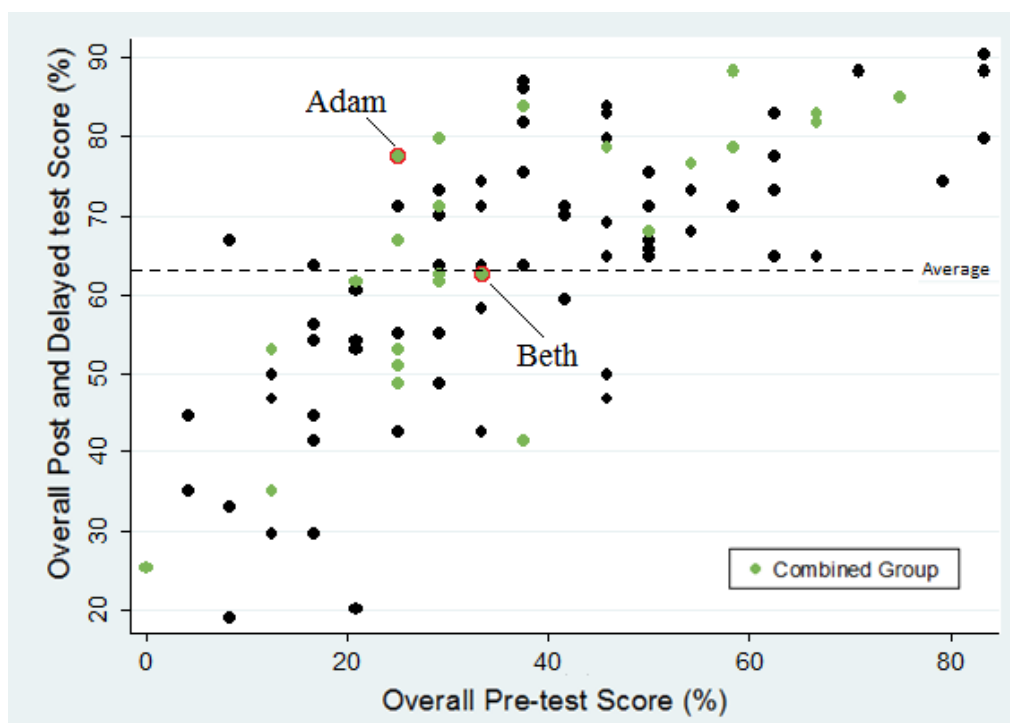


Figure 29: While Adam and Beth had similar pre-test scores, they had different overall post and delayed test scores.

### 6.3 Adam: High Learning Gains

This section presents a case study of Adam from the Combined group who performed well on formal assessments relative to other students, particularly those who began with similar

pre-test scores. First is a description of Adam’s background information and a discussion of his understanding of the CVS at the start of the study. Then, we consider what Adam learned about the CVS, while using the study strategies, based on his performance on the post and delayed tests. Next are separate examinations of Adam’s self-explanations and drawings, with attempts to characterize how meaningful each were in terms of learning about the CVS. Throughout this section, we will consider how Adam seemed to be using the study strategies, which characteristics of his self-explanations and drawings were likely to help him learn about the CVS, and what his self-explanations and drawings can tell us about his developing understanding of the system.

### 6.3.1 Adam’s background

Adam identified himself as an Asian male on the screening assessment. When surveyed about his learning style, he agreed that he was a visual learner and read/write learner, and disagreed that he was an auditory and kinesthetic learner. At the time of this study, Adam was 20 years old, a senior in college, and an electrical engineering and computer science major. The last biology course he reported taking was Advanced Placement (AP) Biology, which he passed over two years prior to his participation in this study. He had never taken a college-level biology course before. Overall, Adam was assumed to not have a particularly strong background in biology but was familiar with the CVS from his high school biology course.

### 6.3.2 Adam before studying (Pre-test results)

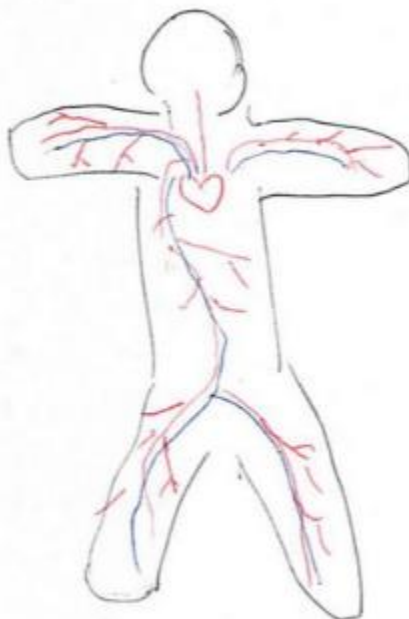
Adam’s SPF sub-scores were below average on the pre-test (Table 21). He responded to 11 of the 18 free-response questions with “I don’t know” or a phrase similar in meaning. Yet, based on the rest of his responses, it was clear that Adam did have some understanding of the CVS, such as the fact that blood is oxygenated in the lungs (something that 30% of the other participants did indicate on the pre-test). Adam also showed some understanding of blood vessels. He correctly defined the term Artery as a large blood vessel that carries blood away from the heart, and Capillary as small vessels that branch from arteries. Additionally, he mentioned that pressure was involved in keeping blood flowing in one direction.

**Table 21:** Adam’s pre-test scores, along with the average scores for all participants and those only in the Combined group.

	Pre-test scores			
	S	P	F	DMM
<b>Adam</b>	29%	0%	31%	3
<b>Overall Average</b>	38%	29%	38%	3.1
<b>Combined Group</b>	44%	24%	36%	3.3

For the pre-test drawing task, Adam generated a very basic drawing at a level of detail that many other participants also generated (Figure 30). One key feature of his drawing was the

utilization of the red and blue colors to represent different blood vessels. In biology text books, oxygenated blood is often represented in red and deoxygenated blood in blue, an exaggerated depiction of the bright red and dark red hues of oxygenated and deoxygenated blood found in humans. Adam may or may not have been aware of this color scheme but, based on his use of these two colors in his drawing and his definition of Artery, it is reasonable to assume that he was aware that two sets of blood vessels exist in the CVS and (at the least) that one carries blood away from the heart. His pre-test drawing was coded as a Single loop without lungs model (3 points) because the two colors suggest that blood travels away from the heart through one type of vessel and back towards the heart through the other.



**Figure 30:** Adam's pre-test drawing of the CVS.

### 6.3.3 Adam after studying (Post and delayed test results)

---

Adam scored higher on all measures of the post and delayed tests compared to the average participant (Table 22). His average SPF sub-scores across the two tests were 80% for structures, 75% for pathways, and 77% for functions. Furthermore, his drawing mental model began as a Single loop without lungs (3 points) on the pre-test and improved to a Double loop-1 (5 points) on the post-test and persisted to the delayed-test (Figure 31). Finally, Adam labeled 74% of the components on the figure interpretation task correctly, and drew the entire path of blood flow correctly.

**Table 22:** Adam’s averaged post and delayed test scores, and average scores for all participants and those only in the Combined group

	<b>S</b>	<b>P</b>	<b>F</b>	<b>Post-test DMM score</b>	<b>Delayed-test DMM score</b>	<b>Labels</b>	<b>Blood Flow</b>
<b>Adam</b>	80%	75%	77%	5	5	74%	10
<b>Overall Average</b>	74%	64%	56%	4.9	4.7	65%	7.5
<b>Combined Group</b>	77%	67%	59%	5.2	5	64%	7.7

Based on his performance on the post and delayed tests, Adam appeared to have learned a lot about the CVS after using self-explanation and drawing to study the CVS text. However, he did not fully and accurately answer all questions on each test and forgot the names of a few terms. On both tests, he did not know what a septum was and this missing piece of information was the main cause of his imperfect assessment scores. On both tests, he could not appropriately define Septum, nor describe the consequences of having a hole in it. On the figure interpretation task (only at the end of the delayed-test), having also forgotten what ventricles were, he mislabeled both of the items pointing to the left and right ventricles as “septum.” Aside from these responses, the rest of Adam’s assessment responses were mostly accurate.

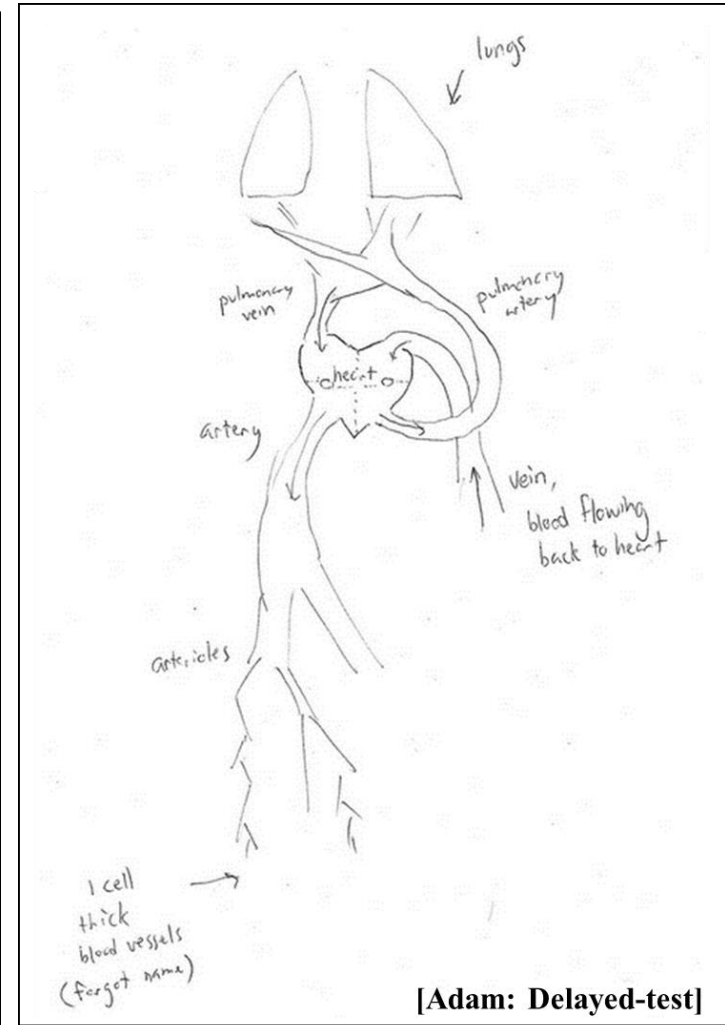
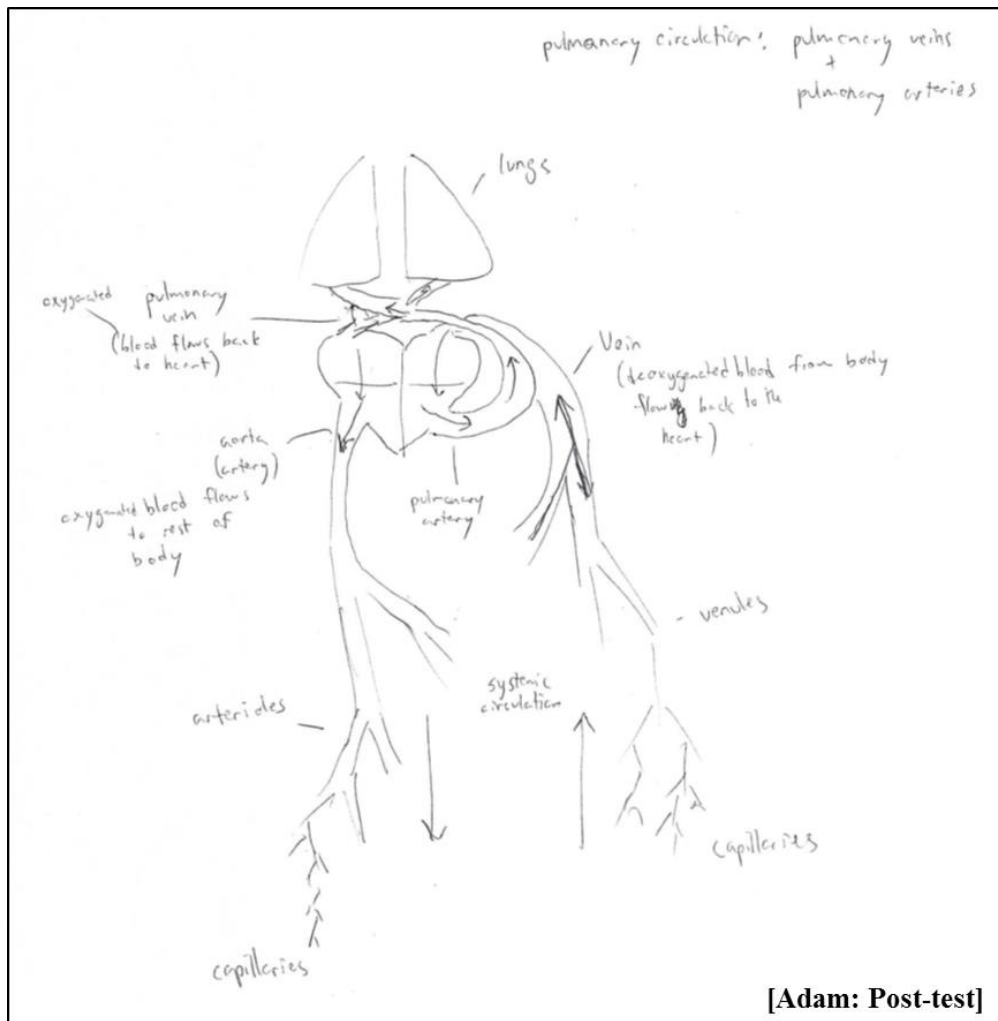


Figure 31: Adam's post and delayed test drawings of the CVS.

#### 6.3.4 Adam's opinions of the study strategies

---

Before examining specific self-explanations and drawings that Adam generated during his study session, let us first consider his opinions about these strategies after he used them in order to inform us of how he may have approached using each strategy to study the CVS text. When Adam finished studying the CVS booklet, he was asked for his opinion of the two study strategies. In response to the question “What’d you think of those two learning strategies?” he responded, “I feel it’s pretty specific to graphic or visual topics. Um, it’s really helpful and efficient for topics like biology and just anything that requires visualization. [...] I feel it was really useful for learning the heart and cardiovascular system.” It is apparent from this quote that Adam’s focus was on the drawing strategy. He suggests that drawing is useful for learning about topics that are visual in nature, including the CVS.

When pressed again for an opinion about both strategies, Adam explained,

I feel like the drawing was more useful and I think partially the reason for that is I’m a very writing-type of person. When I write or draw things, it’s a lot better for me. When I talk and explain things, I think I’m more focused on explaining or the actual speaking part, so I’m not remembering as well. It kind of goes the same when I’m reading something. When I read out loud, it’s less useful than reading silently is. I’m more focused on the pronunciation, or that detracts from memorizing.

Adam seems to believe that self-explaining aloud detract from memorization because his attention is focused on the aesthetics of those activities (i.e., speaking and pronunciation). However, he does not mention being similarly detracted when drawing, even though the products of that strategy have equally evident aesthetics, e.g., clarity and resemblance. Perhaps Adam considers himself a good artist and/or a poor speaker, leading him to pay more attention to how he self-explains but, as he said, in a detracting way.

Given his negative opinions about using self-explanation as a study strategy, Adam was asked if he thought there was anything good about self-explaining. He responded, “It kind of forces you to make connections to your previous knowledge. The more connections you make with something, the more you’re likely to remember it.” This statement suggests that Adam was actually using the self-explanation strategy to make connections, even feeling forced to make these connections. Although he thought the self-explanation strategy was less useful than drawing, he may have been trying to use it effectively.

Adam was finally asked if there was anything *not* good about drawing. He responded, “for smaller concepts [...] it’s sometimes repetitive, or it takes a long time to draw. It does take up some time. There may be other ways that are for memorizing.” From our data, drawing did take significantly more time to complete than self-explaining, but it may be this investment in time to generate these drawings and/or to draw the same concepts multiple times that helps a learner solidify an understanding of the information. The fact that Adam felt that drawing took up a lot of time suggests that he invested some time in his drawings while he studied the CVS text.

Clearly, to Adam, a study strategy that helps him memorize the best or the most, in the least amount of time, is the best strategy for him. He favored the drawing strategy even though it took up more time and, although he found it somewhat detracting, he did seem to understand the theoretical benefits of the self-explanation strategy. While his comments are not an indication that Adam used the self-explanation and drawing strategies to the best of his understanding and

ability, it does imply that he discovered these benefits in each by effectively using them to some degree.

#### 6.3.5 Adam's Self-explanations

---

Considering that Adam scored relatively high on the post and delayed tests, he was expected to have generated more than the average number of self-explanations. Based on his self-explanations of passages 2 and 3, he generated 20 utterances; 3 were goal-driven self-explanations, and 1 was elaborative. Compared to the average 19.2 utterances, with 3.7 being goal-driven self-explanations and 3.3 being elaborative, Adam's numbers were lower than what was predicted by the correlation between self-explanations and assessment scores found in Part 2 of this dissertation.

Considering his verbal data from all 10 passages, Adam still did not seem to generate a particularly large number of self-explanations. From Passages 2 through 11, he generated 88 utterances, 18 goal-driven self-explanations and 10 elaborative self-explanations. In those self-explanations, he mentioned structures 6 times, pathways 5 times, and functions 20 times. He also made one incorrect inference in his explanations. The remainder of his utterances were restatements of the information presented in the text, containing no inferences.

The quality of Adam's self-explanations were consistently high, based on the criteria proposed by Brewer, Chinn, & Samarapungavan (1998) for the evaluation of explanations in science. Adam's self-explanations were regularly accurate, consistent with prior information presented in the booklet, simple, and precise. Furthermore, the purposes his self-explanations served were often to fill gaps in the text, propose specific purposes for pieces of the CVS, and expand the scope of the passage to include a larger picture of the system. Some of his self-explanations accomplished all three of these purposes simultaneously, but the following examples will focus on each of these separately. In the next sections, we discuss characteristics that Adam's self-explanations did *not* have in order to provide a contrast for the subsequent descriptions of characteristics that his high-quality self-explanations *did* have.

##### *a. No explicit effort to memorize*

When asked for his opinions about the study strategies, it was clear that memorizing was a priority feature that Adam used to judge a study strategy's value. However, little evidence of this was found in his verbal protocol. Perhaps his characterization of "memorizing" was more sophisticated than other participants' whose attempts to memorize information were explicit in their verbalizations. Some participants repeated words, phrases, or entire sentences, in succession. Some made word associations in the form of alliterations or mnemonic devices. These statements were likely to help little with understanding. In addition, some learners prefaced these associations with phrases like, "you can remember this by..." or "to help remember this..." Adam did none of these things. Instead, all of Adam's utterances seemed to focus more on understanding the content and less on memorizing pieces of it.

##### *b. An example low-quality self-explanation*

Next, let us consider one of the few self-explanations that Adam generated that was not very meaningful. He generated this self-explanation while studying Passage 3, which describes

all of valves in the heart and the general direction of blood flow between the heart, lungs, and body. After reading through the passage, Adam explained:

Utterance: So the heart has many valves that ensure that the direction of blood flow is correct

A.SE1: because it would be very bad if blood flowed the wrong way or it just wouldn't have a very efficient pump.

The line preceding A.SE1 was not coded as a self-explanation because it did not introduce information beyond what was stated in the text. A.SE1, on the other hand, was coded as a goal-driven self-explanation that mentioned a function because a purpose for the valves, direction of blood flow, or pumping of the heart is vaguely implied. One may argue that this self-explanation is meaningful because Adam is considering the consequences of blood flowing the wrong direction, which is never addressed in the CVS text. Technically, it fulfills some of the criteria listed in Brewer et al. (1998): it is accurate, consistent with other information, and certainly simple. However, stating that the consequences would be “bad” and or result in inefficiency does not fill this gap in the text in a way that would help Adam understand the function of valves. The lack of detail makes this explanation imprecise and also does not widen the scope of the topic. Adam could make analogous statements about any other function described in the text: if a component did not do what it was described to do in the text, the result would be something bad or inefficient.

A more meaningful self-explanation about this sentence would have mentioned the flow of oxygenated and deoxygenated blood. If blood did not flow in only one direction, it would be difficult for oxygenated blood to return back to the heart and also be pumped to body cells, and for deoxygenated blood to return back to the heart and also be pumped to the lungs. Given the information presented in Passage 2 and the self-explanations Adam generated for that previous passage, these were inferences that he could have made.

*c. Fill gaps in the text*

I imagine that Adam's focus on understanding (in contrast to memorizing) is what made his self-explanations sound consistently meaningful. We now consider some of the high-quality self-explanations that were more frequent in Adam's verbal data. These self-explanations were more specific to the CVS content and useful for understanding the system as a whole. Again, these high quality self-explanations often filled gaps in the text, described specific purposes or functions of components, or described how smaller components worked as parts of the larger system.

The following self-explanation is an example of one that most clearly shows Adam filling a gap in the text. While reading most of Passage 4 about how ventricles contract and force blood out through the arteries, Adam did not generate any inferences and only restated the information in his own words. However, after reading and summarizing the final sentence of the passage, Adam generated a self-explanation:

Passage 4: In another phase, called diastole, the ventricles relax, and the blood flows in from the atria.



Utterance: and then when the ventricles relax in the second phase which is called diastole,

Utterance: the blood will flow from the atrium which is the top section into the ventricles, so the lower sections.

A.SE2: And also guessing the blood will flow from the outside vessels into the atrium in that phase.

Note that Passage 4 only describes what happens to the ventricles during systole and diastole, and although the atria are mentioned, they are largely ignored in this passage. In fact, atria and ventricles both contract, separately,<sup>6</sup> but this is never mention in the text.

Perhaps, Adam noticed that Passage 4 did not explain what happened to the atria during these phases and proceeded to contemplate their role during these contraction phases. Given that the passage does not explain that the atria contract, Adam may have assumed that they contract and relax simultaneously with the ventricles. By explaining that blood will flow into the atria (A.SE2), Adam is completing the larger picture of the CVS by imagining blood filling all of the chambers of the heart rather than only the bottom half. This may have helped him to understand the dynamics of heart contractions and how blood flows through all of the chambers

*d. Infer specific purposes*

Some of Adam's self-explanations expanded the scope of the passage by inferring a specific purpose for a structure, pathway, or function being described. For example, Passage 2 explains how blood from the right side of the heart is pumped to the lungs and blood from the left is pumped to other parts of the body. Note that descriptions of the exchange of oxygen and waste at the lungs and body cells does not appear until Passage 3. Adam explained the following as he read Passage 2:

Utterance: And the right side pumps to the lungs

Utterance: and the left side pumps to other parts of the body

A.SE3: and I can imagine that's because the blood needs to get to the lungs or needs to be refreshed with oxygen

A.SE4: and the other side is for delivering that increased nutrients to other parts of the body

In A.SE3 and A.SE4, Adam is inferring that there must be a reason that blood is pumped to the lungs and to the body. While the content of the self-explanations are not particularly insightful because many participants knew that blood is oxygenated in the lungs, making these types of connections are likely to be useful in developing a mental model of this system. These self-explanations definitely expand the scope of the passage about structures and pathways to include information about function, or purpose. Perhaps this self-explanation allowed Adam to more easily associate each side of the heart with oxygenated and deoxygenated blood.

---

<sup>6</sup> As the heart relaxes, pressure inside of it reduces and blood is drawn in from veins, filling both the ventricles and atria. Once passively filled in this manner, the atria contract to force approximately 30% more blood volume into the ventricles. The ventricles then contract and pump their contents into arteries, as the atria relax and begin their passive filling. Overall, the atria and ventricles share some overlap in their contracted and relaxed states, but the overlap is staggered.

e. *Expand the scope of the text*

The next set of example self-explanations are all taken from Adam studying Passage 6, which describes the structure of blood vessels. Adam generates self-explanations about each type of blood vessel by considering how it fits into the larger scope of the CVS. Expanding the scope of the information presented in the passage was something that Adam did regularly with his self-explanations. The passage is comprised of three sentences that each describe the structure of arteries, capillaries, and veins, in that order. We will consider each of these sentences and the self-explanations that Adam generated after reading each of them.

After reading about arteries, Adam explained why they might be the thickest of the blood vessels.

Passage 6: The thick walls of the arteries have three layers, an inner endothelium, a middle layer of smooth muscle, and an outer layer of connective tissue.

Utterance: So structures of blood vessels are that arteries are the biggest vessels, will be the thickest, with three layers.

A.SE5: And uh, it's also because they're containing most of the blood

A.SE6: and it's also how it handles a lot more pressure because of all that blood

A.SE7: and the speed at which it's getting pumped through.

Notice that the text focuses strictly on the structure of blood vessels, and does not mention why the vessels might have these structures. Nevertheless, Adam was able to make inferences about why arteries are the thickest by considering how they are positioned in relation to the rest of the CVS. He infers that arteries contain “most of the blood” and must withstand the pressure and speed with which blood is pumped through them. It is unclear what he means by and why he believes that arteries hold most of the blood, but arteries are the largest and most elastic blood vessels and do contain more blood than the capillaries and veins.

What is clearer is that Adam is referring to the fact that arteries branch out directly from the ventricles of the heart, where blood is pumped out and there is high pressure. This is assumed by his use of words about high pressure (A.SE6), blood being pumped (A.SE7), and speed (A.SE7), and the fact that Adam correctly defined Artery on the pre-test. There were no assessment questions that asked learners why the arteries were the thickest blood vessel, so it is unknown what proportion of the other students understood this concept. However, anecdotally, very few students attempted to self-explain why arteries are the thickest type of blood vessel while studying this passage, and even fewer did so correctly. Without considering that arteries are connected to the heart, where blood is pumped out, a learner would not be able to understand why the arteries are the thickest blood vessel. The fact that Adam made this connection between the purpose of the thick walls and how arteries are positioned within the larger system suggests that he understood the structure of the arteries and what purpose arteries serve in the CVS.

The next line in Passage 6 was about the structure of capillaries. After reading it, Adam generated an inference about why the capillaries might have thin walls:

Passage 6: Capillary walls are only one cell thick.

Utterance: And I'm guessing the reason for it only being one cell

A.SE8: is that the nutrients can be effectively – or that the blood vessels can effectively provide the nutrients or the oxygen to neighboring cells.

From previous passages and Adam's self-explanations, it was clear that he knew nutrients and oxygen are exchanged with waste in the capillaries, and as with his self-explanations about arteries, he considered the purpose of the structure of these vessels. In A.SE8, he infers that the capillary wall's structure allows for this exchange to occur. Less than 65% of all participants were able to make this connection, based on responses to the question about why capillary walls are one cell thick on the post and delayed tests.

While Adam may not have been necessarily thinking about all capillaries in the whole CVS (he only refers to the capillaries that provide nutrients to cells and not in ones that acquire nutrients in the lungs) he *was* considering their role within the system, on a slightly larger scale that included the body cells surrounding the capillaries. Without considering that capillaries are the site of gas exchange, a learner would not be able to understand why their walls are only one cell thick. This self-explanation could help Adam to understand and link together the structure of capillary walls and their role in distributing nutrients in the CVS.

When reading about the last type of blood vessel, the veins, Adam again made inferences about the purpose of its structure by considering how the veins are part of the larger system.

Passage 6: The walls of veins, which are also composed of three layers, are thinner and less muscular than those of arteries.

A.SE9: The veins which are [...] the main vessels coming back to the heart  
Utterance: also have three layers but they are less muscular than the arteries and thinner

A.SE10: and one reason I can think for that is that the cells\* are deoxygenated

A.SE11: and not getting pumped out they're rather just entering the heart

A.SE12: or it's when the heart is getting—it's during the diastole phase where the ventricles relax and the blood comes back in.

A.SE13: So it's not getting forced out it's—it probably has less force on it than the arteries.

\*Note that based on later self-explanations, it was apparent that Adam sometimes used "blood cells" to refer to blood. So, when he states that "the cells are deoxygenated," he is referring to blood.

Adam first considered which direction veins carried blood (which was not restated in this passage; A.SE9), then what kind of blood those veins carried (A.SE10), and where veins were located relative to blood being pumped out of the heart (A.SE11– A.SE13). Surely, he was making sense of the structure of veins by first considering how the veins fit into the larger context of the system. Doing this allowed Adam to realize that there is less pressure in veins than arteries (A.SE13), which was the reason arteries needed thicker structures (A.SE6). Thus, the walls of veins do not need to be as thick because they do not receive blood directly pumped from the heart and actually get blood drawn *from* them by the heart chambers relaxing.

In this excerpt, Adam is making meaningful inferences about veins by considering them in the context of the whole system and deducing the reason veins are not as thick as arteries. He expanded the scope of the text that concentrated on the thickness of veins by also considering the direction of blood flow through the vessels, oxygenation/deoxygenation of blood, phases of muscle contraction, and pressure. Without considering these other components, a learner might not be able to understand why the walls of veins are thinner than those of arteries. These self-

explanations are likely to help Adam better understand the structure of these different types of blood vessels and their purposes in the CVS.

*f. Summary of Adam's self-explanations*

In summary, Adam regularly produced high-quality self-explanations that expanded on the information presented in the text by filling gaps, reasoning about the specific purposes of various components, and considering how specific parts fit within the workings of the larger system. While he ultimately expressed that he was focused on memorizing information, his self-explanations did not reflect this. What they show is that he generated meaningful inferences that could effectively help with understanding the different components of the CVS and how they work together to form an efficient system.

6.3.6 Adam's Drawings

---

Analyses of Adam's drawings for passages 3 and 11 showed that he generated 9 labels and 17 images, compared to the overall average of 14.4 labels and 19.1 images. As with his self-explanations, Adam did not generate more than the average number of images as was predicted by the correlation found in Part 2 of this dissertation between the number of images generated and assessment scores. However, Adam's drawings were similar to his self-explanations in that while they did not contain as many images as expected, they were regularly accurate, consistent, simple, and precise. From passage 2 to 11, Adam generated 42 labels and 64 images of the content presented in the CVS text.

At a glance, Adam's drawings appear minimalistic. Most were simple, organized, and did not cover much of the page. Furthermore, while on his pre-test drawing he used color to distinguish blood vessels from each other, he did not utilize the colored pens as he drew during his study session or on his post or delayed tests. All of the drawings he created after the pre-test were drawn in black ink.

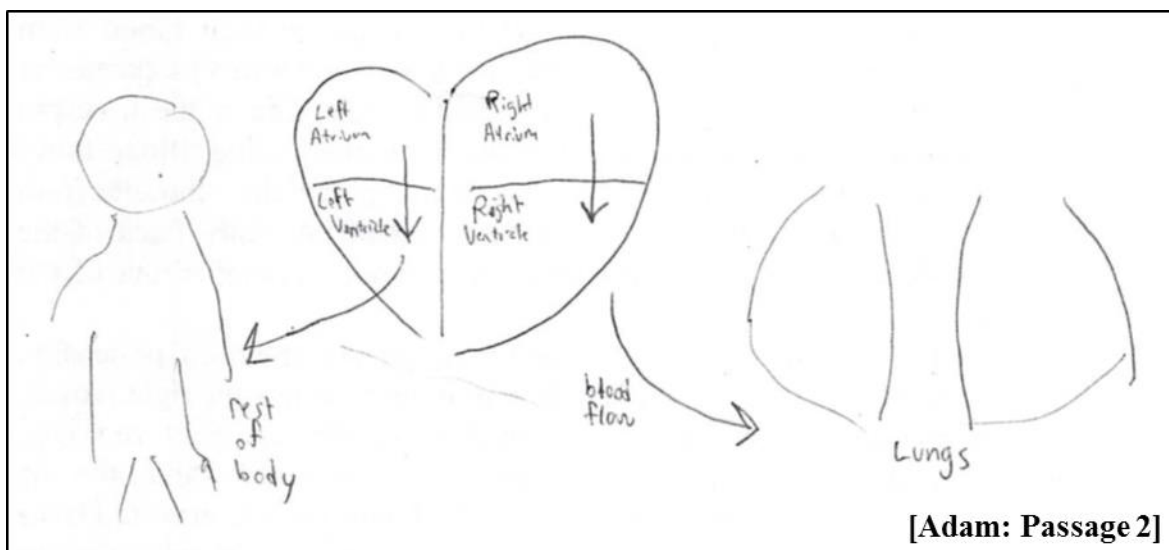
Adam's drawings were noticeably clear and interpretable. One could look at any of Adam's drawings and discern what each of the images and labels represented, and how they connected to each other. This is in contrast to students whose drawings were less orderly, where images sometimes overlapped or were not drawn with complete or distinct boundaries. Furthermore, Adam's drawings were found to be rich with the content provided in the text passages that they represented.

*a. Focus on passage content*

What greatly affects the interpretability of a drawing are decisions a student makes about what to represent and what not to represent, decisions that also greatly affect what content is perceived as the focus of the drawing. Adam's drawings suggest that he consistently tried to concentrate them on the main ideas of each passage.

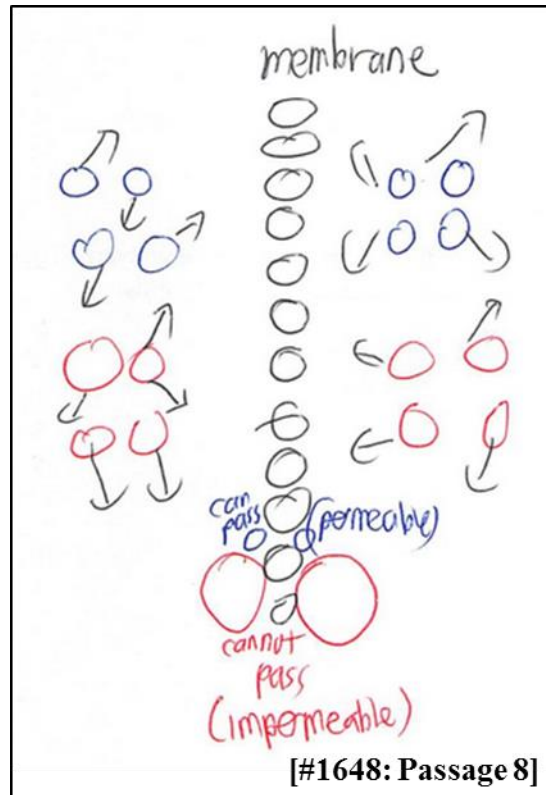
Adam's drawing of Passage 2, which describes the chambers of the heart and where each side pumps is presented in Figure 32. He decided to draw the heart, the lungs, and the body separate from each other, rather than the heart and lungs within the outline of the body, as some other participants did. With this design, the drawing focused on the structure of the heart chambers and where blood flowed to from each side, which were the main ideas in the passage.

Other participants who drew the heart and lungs inside of a body outline had to concern themselves with how the organs would be arranged and how blood would be represented as pumping to the rest of the body from within the body. As a result, some of these drawings would generally appear to represent the flow of blood through the body, rather than the specific structures and pathways emphasized in the passage.



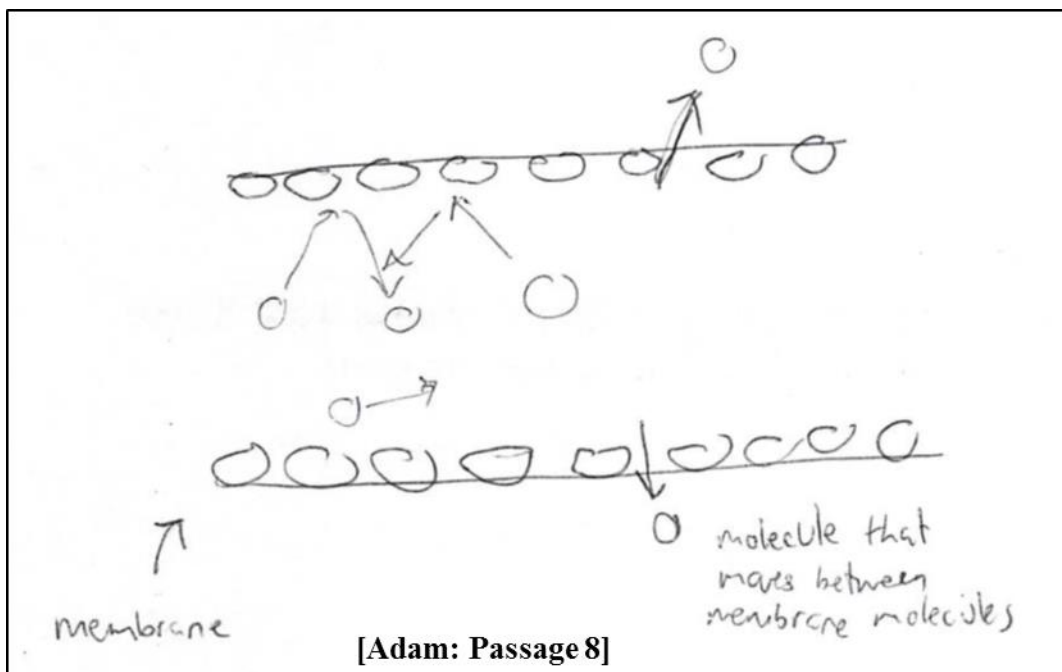
**Figure 32:** Adam's drawing of Passage 2, the general structure of the heart.

In addition to focusing his drawings on the content of each passage, Adam also focused his drawing for Passage 8, the more abstract passage about diffusion through a membrane, on the CVS. The passage describes that (non-specific) molecules move constantly and randomly, and can sometimes pass between the molecules that make up a membrane. Sometimes, when self-explaining this passage, participants would make the connection to the distribution of nutrients in the CVS. When generating a drawing for this passage, however, many participants tended to draw the membrane as described: one random porous wall. This wall usually separated two areas filled with different amounts of unidentified molecules. An example of such a drawing is presented in Figure 33.



**Figure 33:** Participant #1648's drawing of Passage 8, representative of most participants drawing of this passage.

What made Adam's drawing distinct from most others is that he drew two membranes opposite each other, which resemble the walls on the sides of a blood vessel (Figure 34). In addition, he drew molecules between these walls (or within the vessel) either bouncing off the molecules that made up the wall or passing between them. This drawing more closely resembles oxygen or other nutrients being delivered to the cells of the body through the walls of a capillary than most other drawings generated by participants for this passage. If Adam did intend to contextualize his drawing of this passage within the CVS, it is likely to be more effective for understanding the processes associated with the distribution of nutrients in the CVS than drawings like the one shown in Figure 33.

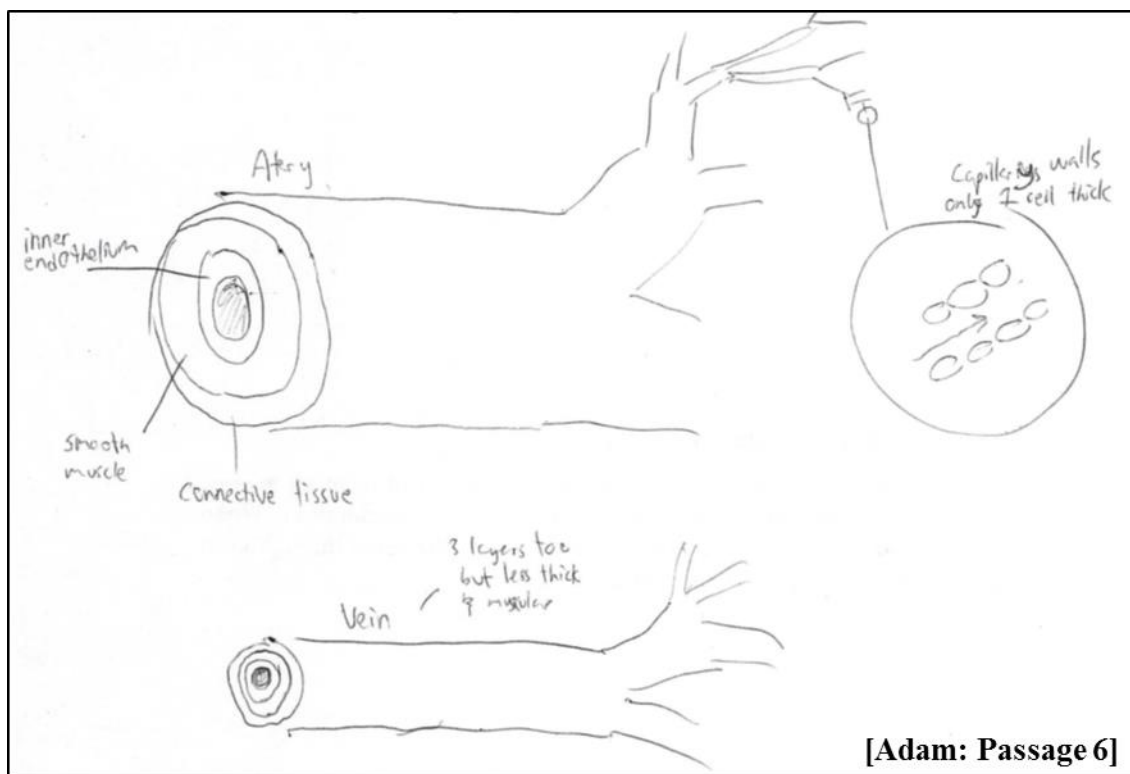


**Figure 34:** Adam's drawing of Passage 8, the process of diffusion through a membrane.

*b. Have a narrow scope*

Closely related to the content focus of a drawing is the scope. If a passage describes one particular component of the CVS, a learner may read this and generate a drawing of that one component alone or of the larger system with an emphasis on that component. In general, Adam's drawings tended to be more in line with the former description, a narrow scope of the content. His drawings seemed to focus on the main ideas of the passages and generally did not observably expand to include more of the CVS than what was necessary to represent these ideas.

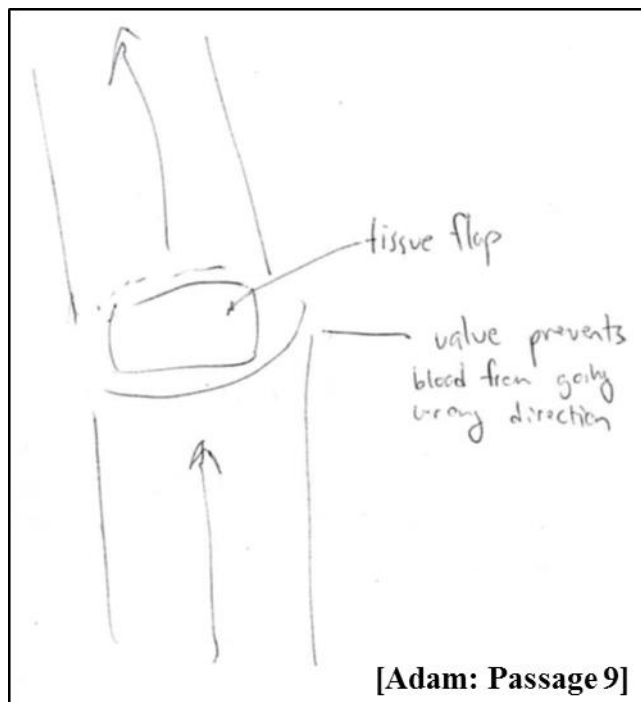
Take, for example, Adam's drawing of Passage 6 (Figure 35), which describes the structure of arteries, capillaries, and veins. This drawing focuses specifically on the content that was presented in the passage, with little added content (other than the inferences about sizes and spatial arrangements necessary to create the drawing). Rather than drawing a complete circle of blood vessels branching from the heart and back to the heart, as he did for the previous passage, Adam focused his drawing on the main idea of the passage: the layers of tissue that make up the walls of the blood vessels. Keep in mind that while self-explaining this passage prior to creating this drawing, Adam inferred that walls of the arteries must be thick in order to accommodate the pressure generated by the ventricles pumping blood directly through them, but he does not make any references to the heart or pressure in his drawing.



**Figure 35:** Adam's drawing of Passage 6, the structure of blood vessels.

Adam's drawing of Passage 9 is another example of a drawing with a narrow scope, concentrated on the content presented in the text. Passage 9 is comprised of two sentences describing that valves prevent blood from moving backward or downward, and allow blood to flow in only one direction through veins. His drawing is of one valve in one blood vessel (Figure 36). In his self-explanation prior to generating this drawing, he inferred that valves must also exist in the arteries, but he does not include both a vein and an artery in his drawing. He also does not make any reference to the valves in the heart. Adam may have been aware that Passage 9 only described the function of valves in veins (because they are described to prevent blood from moving downward, which would eliminate atrioventricular valves in the heart, and valves do not exist in arteries). Thus, his drawing focuses on the function of one valve in one blood vessel (presumably a vein).





**Figure 36:** Adam's drawing of Passage 9, the function of valves.

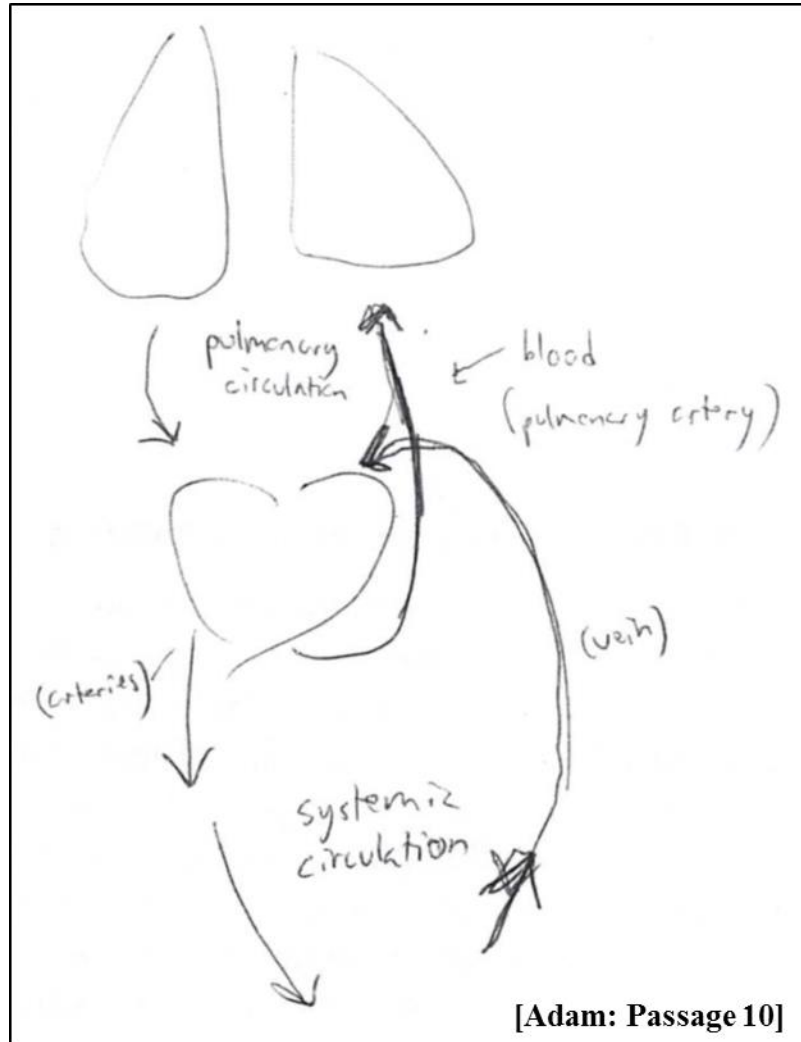
The examples of Adam's drawings presented in this section are representative of the scope of most of the drawings he generated as he studied the CVS booklet. While he tended to expand on the information presented in passages in his self-explanations, the labels and images in Adam's drawings tended to focus on the CVS content from the passages. A narrow scope might also describe the content in Adam's drawing in Figure 34, where he took the abstract and general description of diffusion through a membrane and contextualized it to fit within the scope of the CVS. Overall, the narrow scope of Adam's drawings may have been helpful for him to concentrate on the main ideas in the passages, and generate complete drawings that were clear and focused.

*c. Exhibit some level of abstraction*

Levels of abstraction can vary in drawings when students try to represent a passage that describes an abstract concept, such as the diffusion of molecules across a membrane (Passage 8). It can also vary when students draw a passage that describes concrete concepts. This tended to occur less frequently, probably because individual images often just represented the concrete parts described in the passage. (Although, on a few occasions, a participant only used a label to represent a component, sometimes with a box drawn around the label.) More often, the individual images would take on realistic shapes, but the organization of the images would be abstract. For example, Adam's drawing of Passage 2 (Figure 32) contained realistic components (i.e., the body, the heart, and the lungs each took on somewhat realistic shapes) but they were organized in an abstract manner, where the heart and lungs were drawn next to the figure of the body rather than inside of it.

For another example, see Adam's drawing of Passage 10 about the two primary subsystems of circulation (Figure 37). Images of the lungs and heart were drawn somewhat realistically, but all of the blood vessels that branch and connect the heart to the lungs and the body are collapsed into single arrows. While these arrows indeed have concrete, physical

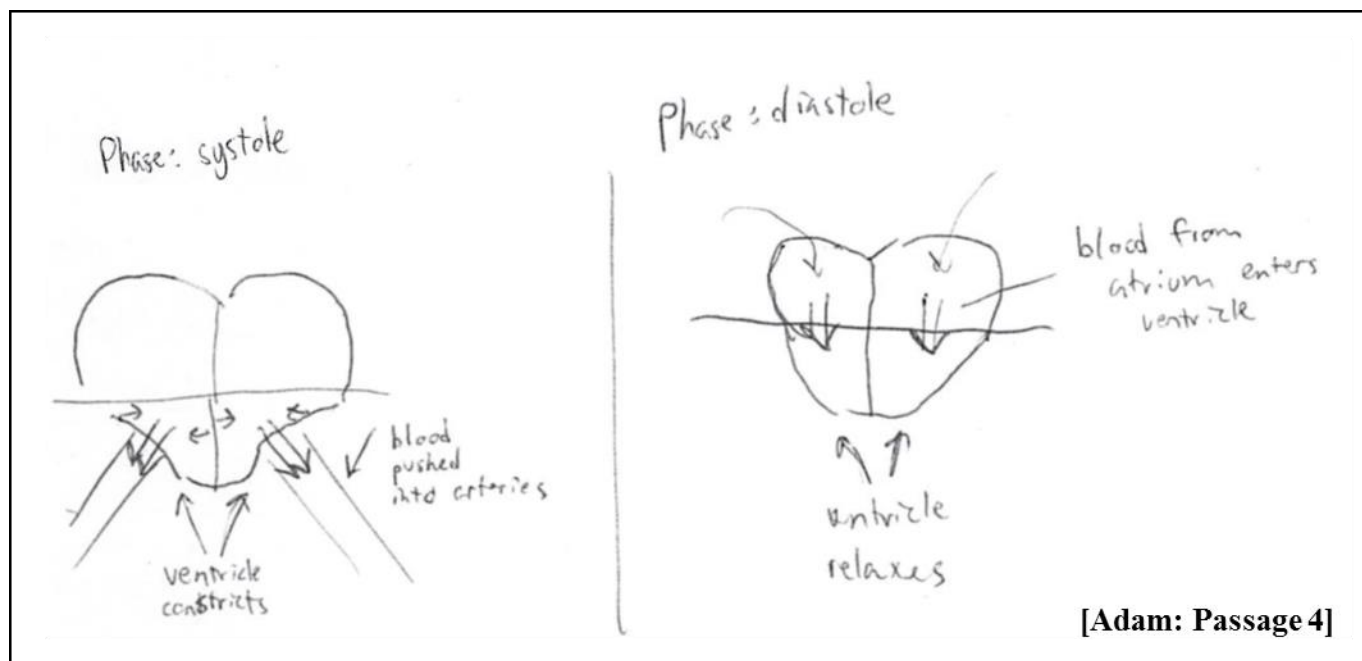
referents, in this drawing, they serve as symbolic representations of blood vessels. This abstraction removes details that are extraneous to understanding the pulmonary and systemic subsystems and focuses the content of this drawing on these two loops. It is clear that the pulmonary subsystem involves the heart and the lungs, and reasonable to assume that the systemic subsystem involves the heart and the body.



**Figure 37:** Adam's drawing of Passage 10, the two primary subsystems of circulation.

*d. Manage motion and time*

Most of the CVS passages and associated drawings have elements of motion in them. A majority of the motion described in the text is related to blood traveling from one place to another. This was represented as arrows in all participants' drawings, not a particularly difficult motion to represent. One passage, however, described a specific motion that was not just blood flow. Passage 4 describes the two phases of contraction of the heart: systole, when the ventricles contract, and diastole, when the ventricles relax. Adam decided to generate two images of hearts for his drawing of this passage (Figure 38).



**Figure 38:** Adam’s drawing of Passage 4, phases of muscle contraction.

In general, this is a good way to represent the two phases of muscle contraction. Trying to include and distinguish the two phases on one image of a heart can get cluttered and confusing. On the left side of his drawing that represents systole, Adam shows the walls of the ventricles constricting, using arrows pointing inward in each ventricle and a label that states “ventricle constricts.” Blood is also drawn as arrows and labeled leaving the heart into arteries. On the right side of his drawing that represents diastole, he does not use arrows to show the walls of the ventricles relaxing, but labels this motion with “ventricle relaxes.” Blood is also drawn and labeled entering the ventricles through the atria, and originally from veins. He could have also included a representation to indicate that the heart cycles between these two phases, but this cyclical process was not described in the text. Although there were few opportunities to represent time while drawing these passages (aside from the movement of blood), Adam did not seem to have any problem with representing temporal sequences.

*e. Summary of Adam’s drawings*

In summary, Adam produced drawings that were easily interpretable. They tended to highlight the main ideas in the text passages, focusing only on the relevant components. His use of some abstraction helped with the interpretability and focus of the drawings by eliminating unnecessary features. Finally, descriptions of motion in the CVS were no challenge for Adam to represent in his drawings. The clarity and quality with which Adam produced his drawings suggest that he had a good grasp of the material.

## 6.4 Beth: Average Learning Gains

This section presents a case study of Beth, also from the Combined group. In this case study, we first present a description of Beth’s background information and discuss her understanding of the CVS at the beginning of the study. Then, based on her responses to the post and delayed tests, we consider what she learned about the CVS while self-explaining and drawing. Finally, we discuss Beth’s self-explanations and then her drawings, characterizing how

useful they could be for learning about the CVS. Throughout this analysis, we will consider how Beth seemed to have been using the study strategies, which characteristics of her self-explanations and drawings were likely to help her learn about the CVS, and what her self-explanations and drawings can tell us about her developing understanding of the system.

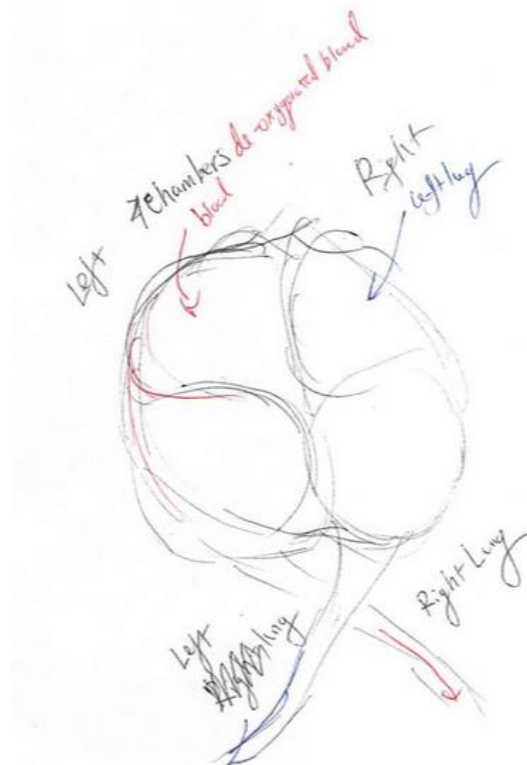
#### 6.4.1 Beth before studying (Pre-test results)

Beth identified as an Asian female on the screening assessment. When surveyed about her learning style, she strongly agreed that she was an auditory learner and a read/write learner, neither agreed nor disagreed that she was a visual learner, and strongly disagreed that she was kinesthetic learner. At the time of the study, Beth was 18 years old, in her sophomore year of college, and majoring in business. The last biology course she reported taking was MCB61: Brain, Mind, and Behavior, two semesters prior to this study. MCB61 is a college-level biology course in neuroscience, concentrating on the brain and nervous system, memory, language, emotion, and the mind. The details of the CVS are not likely to have been a part of the course material. In addition to this course, Beth reported taking AP Biology in high school and passing the course. Overall, Beth was assumed to have some background in biology and may have been a little more familiar with the CVS than Adam since she was younger and completed AP Biology more recently. These differences were detectable on the pre-test, as Beth scored slightly higher than Adam on structures and pathways sub-scores, and on the drawing task (Table 23).

**Table 23:** Beth and Adam’s pre-test scores, along with the average scores for all participants and those only in the Combined group.

	Pre-test scores			
	S	P	F	DMM
<b>Beth</b>	<b>43%</b>	<b>25%</b>	<b>31%</b>	<b>4</b>
<b>Adam</b>	29%	0%	31%	3
<b>Overall Average</b>	38%	29%	38%	3.1
<b>Combined Group</b>	44%	24%	36%	3.3

Of the 18 questions on the pre-test, Beth only used the response “I don’t know” once. This was when asked to define Septum. She answered 3 questions correctly, 9 questions partially correctly, and 5 questions incorrectly. She had understood of the general structure of the heart, that blood acquires oxygen in the lungs, and the function of capillaries. She seemed to be vaguely aware that two phases or systems are involved in the CVS and that oxygenated blood needs to travel to the brain, but never addressed the rest of the body. It is unclear if she recognized veins and arteries as different types of blood vessels, as she mistakenly defined Artery as atria, identified veins as the vessels that carry oxygenated blood, and described veins to “carry blood north.” Lastly, she answered that the right ventricle pumps blood to the left lung, and included this pathway in her drawing along with the left ventricle pumping to the right lung (Figure 39).



**Figure 39: Beth's pre-test drawing of the CVS.**

Beth chose to use red and blue ink in her pre-test drawing, but the colors did not appear to be used to make any kind of meaningful distinction, other than perhaps blood on one side of the heart versus the other. The left side of the heart is drawn receiving de-oxygenated blood (presumably from the body), and exiting the bottom chamber to go to the right lung. The right side of the heart is illustrated receiving blood from the left lung and also sending blood to the left lung. Assuming that the “deoxygenated blood” written at the top of the drawing applied to both sides of the heart, Beth's pre-test drawing mental model might be described as follows. Blood travels from one side of the heart to the lungs on the opposite side to get oxygen, then delivers oxygen to the body, and returns to the same side of the heart. Based on post and delayed test drawings (Figure 40), it is likely that this is an accurate description of the model she had imagined at the time of the pre-test. For these reasons, Beth's pre-test DMM score it was leniently coded as a Single loop with lungs model (4 points).

Overall, upon entering this study, Beth seemed to have an understanding of a few more details of the CVS than Adam, such as the structure of the heart and perhaps some concept of a separate loop to the lungs. However, she appeared to have an incomplete and inaccurate mental model of the system. Where Adam may have had gaps in his knowledge of the CVS, Beth may have had inaccurate knowledge.

#### 6.4.2 Beth after studying (Post and delayed test results)

Beth's pre-test scores were slightly higher than Adam's but her post and delayed test scores were very close to the average participant's scores or a little below (Table 24), which were below Adam's scores. Her average SPF-scores across the two tests were 73% for structures, 69% for pathways, and 54% for functions. She labeled 53% of the components on the figure interpretation task correctly and only drew one loop on the figure correctly, well below the

average scores on these tasks.

**Table 24:** Beth and Adam's average post and delayed test scores compared to average scores for all participants and those only in the Combined group.

	Post and Delayed test			Post-test DMM score	Delayed-test DMM score	Labels	Blood Flow
	S	P	F				
<b>Beth</b>	73%	69%	54%	5	4	53%	2
<b>Adam</b>	80%	75%	77%	5	5	74%	10
<b>Overall Average</b>	74%	64%	56%	4.9	4.7	65%	7.5
<b>Combined Group</b>	77%	67%	59%	5.2	5	64%	7.7

Based on her post and delayed test responses, Beth appeared to have learned about some aspects of the CVS by using the self-explanation and drawing strategies to study the CVS text. She showed some evidence of understanding that oxygenated and deoxygenated blood need to be separated, the heart is divided into chambers, there is a distinction between blood vessels, two sub-systems exist in the CVS, and pressure is involved in circulating blood. This list essentially covers all main and general aspects of the CVS, but we find that the details of her understanding of each were often inaccurate or missing from her assessment responses.

An idea that Beth repeats throughout her responses on the post and delayed tests was that oxygenated and deoxygenated blood are prevented from mixing in the CVS. She mentions this idea in her responses to 6 questions that appear on both the post and delayed tests, when the concept is really only relevant to the one question that asks what would happen if there were a hole in the septum. Furthermore, on both the post and delayed tests, Beth correctly described the atria and ventricles as the top and bottom chambers of the heart, respectively. However, on the post test, she specified that the atria hold deoxygenated blood and later that the purpose of the atria in the CVS is to serve as "a holding chamber for blood and makes sure that de-oxygenated and oxygenated blood are separate." This latter phrase about separation was echoed in her response on the delayed test.

Beth correctly defined septum in the definition portion of the free-response questions (i.e., the wall that separates the left from the right side of the heart), but she responded to a later question with, "[a] hole in the septum would mean that the atrium is leaking into the ventricles, and oxygenated and deoxygenated blood are mixing." Also, on the delayed test, she defined a septum as "what separates the atrium and ventricles," and for a question on both tests, she explained that blood from the body needs to go to the heart before it goes to the lungs so that oxygenated and deoxygenated blood do not mix. Based on these responses, it might be interpreted that Beth believes the atria hold deoxygenated blood while the ventricles hold oxygenated blood, but how this fits with the rest of her mental model of the system is unclear.

Beth's pre-test drawing was leniently coded as a Single loop with lungs model (4 points), improved to a Double loop-1 model (5 points) on the post-test, and regressed back to a Single loop with lungs model (4 points) on the delayed-test (Figure 40). Two noticeable items were included in her post-test drawing, but not in her delayed-test drawing. First, the left side of the

heart had a “body” label with a bracket from atrium to ventricle (item 1) and the right side had a “heart/lungs” label also with a bracket from atrium to ventricle (item 2). Second, are images and labels of an artery branching into capillaries drawn coming from the heart, between the two lungs, to the body (item 3). In the delayed-test drawing, however, the blood vessels from the heart are not shown to separate between the body and the lungs, and looks as if the blood goes from the heart, to the lungs, to the body, and then back to the heart.

Based on Beth’s free-response answers and drawings, it seems as if she had a more accurate understanding of the CVS at the post-test than the delayed-test, but was inconsistent with some aspects of the system on each test. Overall, Beth’s understanding of the CVS improved from the pre-test after studying the text, but was inaccurate in several areas.

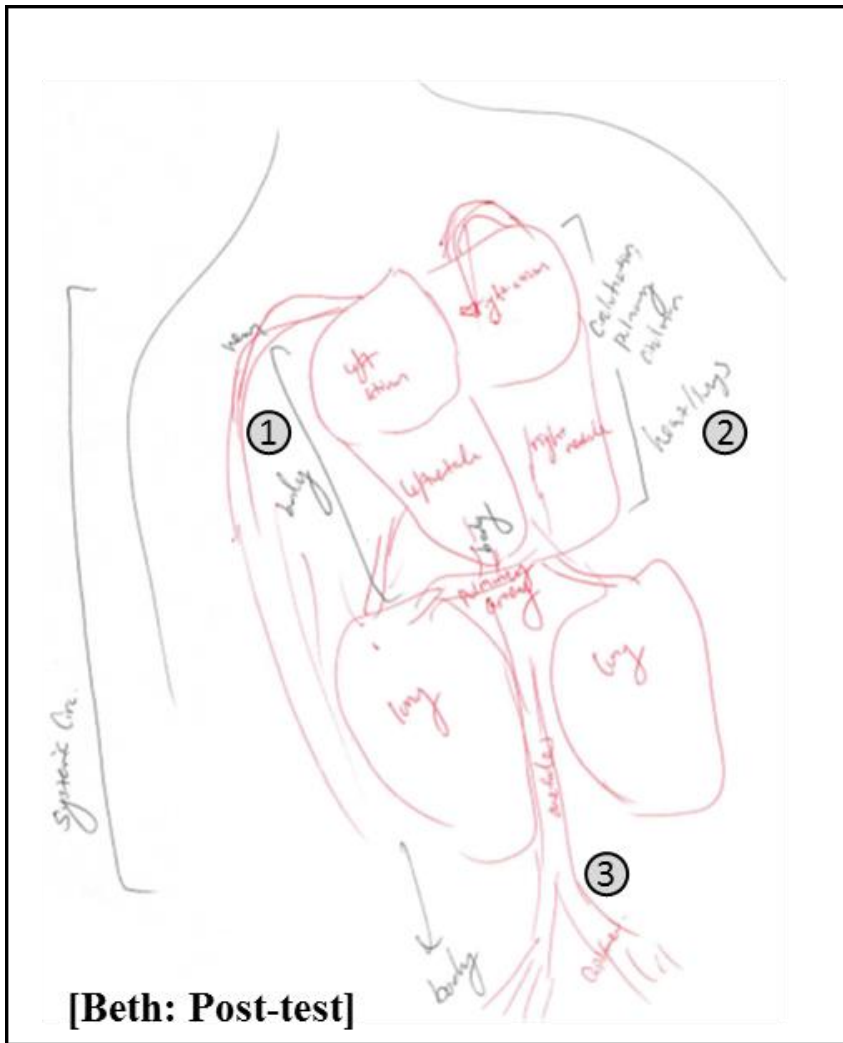


Figure 40: Beth's drawings of the CVS from the post and delayed tests



### 6.4.3 Beth's opinions of the study strategies

---

At the end of her study session, Beth was asked to provide positive and negative comments about self-explaining and drawing as study strategies. Her positive comments about self-explanations were, "it helps you maybe, like, look at concepts better and maybe make connections between different concepts. When you're just reading, you probably won't be able to make the concepts [*sic*] unless you, like, actually say it out loud." While this response is not very informative of how Beth may have been using the self-explanation strategy, it does suggest that she discovered connections while she was self-explaining.

Later, Beth explained that she thought drawing was "easier" than self-explaining because with drawing she could refer back to what she drew, but with self-explaining she could not refer back to what she said, and she often forgot what she said aloud. She had other interesting opinions about how self-explaining was challenging, in the following exchange:

- Beth: If you don't have an understanding of the material before, it's really hard to make connections. Like if you're confused about it or anything, then maybe just talking about it to yourself will make it more confusing.
- Interviewer: Did that happen to you?
- Beth: Yeah, I think one of the passages, I think, about, like – I started getting confused about the left and the right and the aorta. Like, the more I said it to myself, I was, like, "Wait, like, is that right?"

What is interesting is that while Beth initially stated that self-explaining helps to make connections, here, she explains that those connections are hard to make if one does not understand the material first. Perhaps Beth saw self-explanation as a useful strategy for making connections to *other* ideas, outside of the text, only after understanding the connections between the ideas presented in the text. We will see in the next section, times when a lack of understanding of the ideas in the text lead to rather meaningless self-explanations.

Beth's only negative comment about drawing was that it could be difficult for someone to illustrate what they have imagined if that person is not artistic but also immediately stated that the drawings were simple, implying that it requires minimal artistic ability to use this study strategy. Finally, she mentioned that self-explaining before drawing helped "because saying it out loud helps you know what you want to draw, so it gives you the preconceived idea of what you want to draw." This suggests that Beth's self-explanations may have been affected by her anticipation of the drawing task.

Overall, Beth's opinions about the study strategies are not particularly informative about how she may have been using them during her study session. However, they do alert us to difficulties she perceived when self-explaining.

### 6.4.4 Beth's Self-explanations

---

This section will present an overview of Beth's use of the self-explanation strategy. Based on her self-explanations for passages 2 and 3, she generated 18 utterances. Three of those utterances were goal-driven self-explanations and 3 were elaborative. These numbers were near the average for all participants in the study. Surprisingly, however, when considering her verbal data from all 10 passages, Beth generated considerably more self-explanations than Adam: 37 goal-driven self-explanations, 17 elaborative self-explanations. Not only were their self-explanations different in quantity, they were also noticeably different in quality. It is obvious that

the number of self-explanations Beth generated and her formal assessment scores do not align well with the correlation seen between these two variables in Part 2 of this dissertation because she generated well above the average number of self-explanations but only achieved an average score on the post and delayed tests. A closer examination of the content of her self-explanations may help to explain this misalignment.

*a. Explicit efforts to memorize*

Beth did not mention memorizing, as Adam did, when she gave her opinions of the study strategies, but several occasions were identified where she seemed to be using the self-explanation strategy strictly to memorize information. Some had very superficial relationships with the information from the passage and seemed to focus mostly on creating signals for recalling the information from memory, such as stating that “A” comes before “V” in the alphabet to remember that blood flows from Atria then to Ventricles.

Most inferences aimed at strictly memorizing information in this manner are less useful for understanding the CVS as a whole working system. For example, while studying Passage 5 about blood vessels, Beth explained the following to herself:

Utterance: The arterioles um, then again, um split into even smaller vessels called capillaries

B.SE1: and you can just think about it like caterpillars are super tiny so they're the smaller version of even the arterioles

Since the information about caterpillars was not described in the text, this utterance was necessarily coded as an inference. It is an elaboration on the structure of capillaries, introducing new information that they are small like caterpillars. Obviously, this self-explanation will neither help Beth understand the pathway connected to capillaries nor their function, but it may help her to remember that they are very small structures.

Other inferences that Beth generated that were focused on memorization may have been more meaningful. For example, after reading Passage 4, which explains the phases of muscle contraction in the heart, she explained:

Utterance: You can remember the difference between the systole and the diastole um which is when the ventricles relax,

B.SE2: is when someone dilates your eyes that's when your pupils get bigger

Utterance: so when di – through the diastole movement your heart is relaxing

In this example, Beth used a metaphor to help her remember and/or understand that the term diastole refers to the relaxation of the heart muscles. The information explained in B.SE2 was an elaborative self-explanation because it provided additional details beyond the text, inferring that the function (or behavior) of pupils dilating is similar to the heart relaxing during diastole.<sup>7</sup> Using the definition of prefixes to associate with behaviors may help Beth remember that diastole is a phase of muscle relaxation. While this specific association might not be particularly helpful to Beth in connecting to other parts of the CVS, the fact that functions are often more difficult for learners to understand than structures (and presumably pathways), and that functions are inherently associated with other information in the form of a result or an outcome, makes this self-explanation higher in quality than B.SE1 about caterpillars.

*b. Lack precision and detail: non-explanations*

What was likely the result of Beth's focus on memorization was that Beth's self-explanations were often low-quality (even the ones not explicitly about memorization). Even when the inferences she made were focused on the CVS content, Beth's self-explanations seemed to fall short of the criteria for high-quality explanations described by Brewer et al. (1998). The majority of self-explanations that Beth generated were vague and lacked the necessary detail to make them meaningful. Some were more closely related to the information in the text and more relevant to understanding the CVS than the ones she generated for memorization, but they generally were low in quality.

Consider the following example. While studying Passage 6 about the structure of blood vessels in the CVS, Beth attempted to explain the purpose of each of the 3 layers of tissue that made up arterial walls. In explaining the reason arteries need an inner endothelium, she stated the following:

---

<sup>7</sup> Pupil dilation is actually caused by the *contraction* of dilator muscles, but Beth can still use and may have used this metaphor to think of pupils and ventricles both getting bigger.

Passage 6: The thick walls of the arteries have three layers, an inner endothelium, a middle layer of smooth muscle, and an outer layer of connective tissue.

Utterance: and inside would be closer to the blood

B.SE3: so it wouldn't make sense for the muscle to be in direct contact with the blood so it has an inner endothelium

The additional information that she introduced in this self-explanation was that blood is not intended to make direct contact with muscle. Had Beth provided details for the reason behind this separation, rather than that “it wouldn’t make sense,” this self-explanation may have been more meaningful. However, this was the only statement she made about the purpose of the inner layer.

Another example is from Beth’s self-explanation of Passage 5 about the branching of blood vessels. While self-explaining this passage, Beth made three connections for memorization, but those are omitted from the following excerpt to focus on the other, potentially more meaningful self-explanations that she generated for this passage.

For reference, Passage 5 is replicated below, followed by some of the self-explanations Beth generated while reading it.

Passage 5: **Blood Vessels:** The large, muscular vessels that carry blood away from the heart are called arteries. Blood travels through a network of smaller arteries, which in turn divide and form even smaller vessels called arterioles. The arterioles branch into a fan of tiny vessels called capillaries. De-oxygenated blood flows through capillaries that merge and form larger vessels called venules. Several venules in turn unite to form a vein, a large blood vessel that carries blood to the heart

B.SE5: So since we have so much blood coming in through the heart, we need pathways to be able to move the blood away into the heart.

Utterance: So our network that we have are called arteries,

B.SE6: and obviously we can't have the same large blood vessel traveling throughout the body because that would create no pressure

[...]

B.SE7: So deoxygenated blood flows through the capillaries because they're the smallest, um, blood vessels,

Utterance: and they merge to form um larger vessels called venules

Utterance: and several venules form the vein

Utterance: and that is the method to be transported back to the heart

B.SE8: because um once the blood goes through that entire process from large to small we know that it is time for it to go back to the heart

Notice that many of the statements that Beth made were points that were not addressed in the passage. She appeared to consistently try to extend the information from the text but fell short because of the lack of details. The reasons she provided for various things happening were not really reasons at all. In B.SE5, Beth confusingly explained that the reason we have pathways that move blood “away into the heart” is because a lot of blood moves through the heart. There is a sense that Beth is trying to explain a goal, reason, or purpose for blood moving away or into the heart here, but there is also a sense of circular reasoning. Unless Beth actually meant something

different than what she actually verbalized, this self-explanation does not promote learning. Similarly incomplete explanations were made in B.SE6 – B.SE8.

In B.SE6, Beth explained that having “the same large blood vessel traveling throughout the body” would result in no pressure. Presuming that she meant if blood vessels did not get smaller as they spread throughout the body, it is still unclear how this would eliminate pressure. In B.SE7, she explains the reason deoxygenated blood flows through capillaries is because they are the smallest type of blood vessel. It appears that Beth simply took the line from the text, Deoxygenated blood flows through capillaries, continued the phrase with the conjunction “because” and proceeded to describe the structure of capillaries. This ad lib explanation may have just been a compilation of words that sounded like an explanation. She used the word “because” again in B.SE8, technically making the sentence sound like an explanation, but essentially saying that blood goes back to the heart because “we know it is time.” This lack of a true reason makes this statement rather meaningless for learning about the CVS.

*c. Shift scope*

In contrast to Adam’s self-explanations which frequently expanded the scope of the information presented in passages to contextualize components within the larger CVS, Beth’s often did not piece components together or connected to information that was not specifically useful for understanding the system. While studying Passage 4 about the phases of muscle contraction, Beth explained the following to herself:

B.SE4: Um so the contractions is what we hear, what we feel as well when we put our hands on our chest.

While this self-explanation is certainly more connected to the CVS than B.SE1 about caterpillars, how useful is it for learning about the CVS? It expands the scope of the topic to include the senses of sound and feeling,<sup>8</sup> but the inference is very concentrated on the contraction of the ventricles. It does not necessarily fill a gap in the information presented in the text because it does not connect to other information specifically regarding this system. It describes effects of the heart contracting but not one that is related to the functioning of the CVS. And it does consider cardiac muscle contractions at a broader level, from the perspective outside of the body, but also not as part of this specific system. This inference added information to the passage, but tangential information. This self-explanation seems more to have shifted the scope of the passage, rather than expanded it.

*d. Incorrect self-explanations*

While studying Passage 11, the final passage in the CVS booklet, Beth generated many incorrect explanations. Her tone and pacing did not show signs of fatigue but, even so, Beth may have put less effort into these self-explanations than she did into earlier passages. In each of the self-explanations in the following examples, Beth said something that was either an incorrect assumption, contradictory to information presented in the passage or an earlier passage, and/or contradictory to something Beth had explained earlier.

---

<sup>8</sup> Heartbeat sounds are actually created by the abrupt closing of the atrioventricular and semi-lunar valves just after the ventricles contract and relax, respectively. While Beth’s inference is not entirely accurate, the relationships between heart sounds and contractions are close enough that they do not take away from the connection she is making.

Parts of Passage 11 are presented below in consecutive order but separated by Beth's utterances. The passage describes the details of the primary sub-systems of circulation and generally summarizes the path that blood takes through the system.

Passage 11: Oxygenated blood then flows into venules, which merge into the pulmonary veins that lead to the left atrium of the heart. The pulmonary veins are the only veins that carry oxygenated blood.

Utterance: Um... So oxygenated blood flows into the lower -- um, to the veins and it merges into the pulmonary veins so that...

B.SE9: Yeah so the pulmonary veins are the veins that carry blood away from the heart to the other parts of the body

After reading that pulmonary veins lead *to* the heart, Beth stated that the pulmonary veins carry blood *from* the heart to the body (B.SE9). She may have only attended to the words describing that the vessels carry oxygenated blood, and mentally followed the pathway from the systemic side of circulation, which would bring oxygenated blood to the body's cells. However, she disregarded the fact that she had previously read that veins carry blood towards the heart, and explained that pulmonary circulation referred to circulation "involved with the heart and the lungs."

Passage 11: Systemic Circulation is the movement of blood to all parts of the body, except the lungs, and then back to the heart. Blood from the left ventricle enters systemic circulation through the aorta.

Utterance: and this is in relation to systemic circulation because the system is what moves blood through to other parts of the body

Utterance: -- while pulmonary circulation only talks about the heart and -- so that's why it um it excludes the lungs

Utterance: and the systemic circulation use the deoxygenated [*sic*] back to the heart so it can be re-oxygenated.

B.SE10: So the left is always dealing with deoxygenated blood.

B.SE11: and it enters through the aorta.

In B.SE10, Beth makes an incorrect self-explanation that would greatly affect her model of the CVS. If she ends her study session believing that the left side of the heart contains deoxygenated blood, then the next reasonable location for it to travel to is to the lungs to get oxygen (but the left side of the heart actually contains oxygenated blood that pumps to the body). This would be consistent with the drawing from Beth's delayed-test where each ventricle was drawn connected to a lung. However, on her post-test, she answered that the right ventricle pumps blood to the lungs and that blood goes to the left ventricle after it is oxygenated.

It is also possible that after reading the first sentence of this text excerpt, stating that systemic circulation moves blood to the body then back to the lungs, Beth assumed that the following sentence would describe the continuation of that pathway (instead of back to where systemic circulation began from the heart). She may have only attended to "left ventricle" and "through the aorta" in the sentence and assumed those referred to where blood went next.

These incorrect self-explanations are difficult to map onto Beth's responses on the post and delayed tests because her responses were inconsistent within and across the tests. Furthermore, it is uncertain which of and to what extent these incorrect self-explanations truly reflected Beth's understanding of the information versus simple mis-wording or mis-phrasing.

However, overall, the fact that Beth generated many more of these incorrect self-explanations suggests that she was, to some extent, developing an inaccurate understanding of the CVS as she self-explained.

*f. Summary of Beth's self-explanations*

In summary, Beth generated many self-explanations but many of them were not high-quality. While using the self-explanation strategy to learn, she tended to explicitly focus on memorization, not include precise details, not expand the scope of the topic to include the larger system, and generate incorrect statements about the information she read. Some of her self-explanations were inconsistent, or added peripheral or tangential details. Especially in comparison to Adam's high-quality self-explanations, Beth's self-explanations were often not scientifically meaningful. She seemed to have the intention of expanding upon the information presented in the text and forming connections with her prior knowledge, but some or parts of her self-explanations did not have much value for learning about the whole CVS.

#### 6.4.5 Beth's Drawings

Analyses of Beth's drawings for passages 3 and 11 showed that she generated 11 labels and 16 images, compared to the overall average of 14.4 labels and 19.1 images, and Adam's 9 labels and 17 images. From passage 2 to 11, Beth generated 49 labels and 67 images, compared to Adam's 42 labels and 64 images. Overall, her numbers were close to Adam's and a little under the average, both overall and within the Combined group (Table 25).

**Table 25:** Beth's and Adam's booklet drawing scores compared to average scores for all participants and those only in the Combined group.

	Passages 3 and 11		Passages 2 through 11	
	Labels	Images	Labels	Images
<b>Beth</b>	11	16	49	67
<b>Adam</b>	9	17	42	64
<b>Overall Average</b>	14.4	19.1		
<b>Combined Group</b>	12.7	17.6		

Overall, Beth's drawings appeared to reflect her self-explanations in that they were sometimes not well-focused and difficult to interpret. While she did generally represent the information presented in each passage, some of the details were either indistinguishable or incorrect. Especially in contrast to Adam's drawings, Beth's were sometimes unclear, with images overlapping and labels with undistinguished associations.

Again, remember that each participant was asked to draw for him or herself and no other audience. Thus, it is difficult to comment on the meaningfulness of Beth's drawings when parts were found to be uninterpretable because the drawings may have been clear to Beth. Although, given some inconsistency and errors in her drawings, her drawings were not likely to be as useful for learning about the CVS as Adam's drawings.

The quality of Beth's drawings will generally be described in the following sections based on the same criteria used to evaluate her self-explanations: how useful they can be for learning about the individual components of the CVS and how the components work together in a functioning system. We will also consider the same criteria used to describe Adam's drawings: content focus, scope, level of abstraction, and representation of motion.

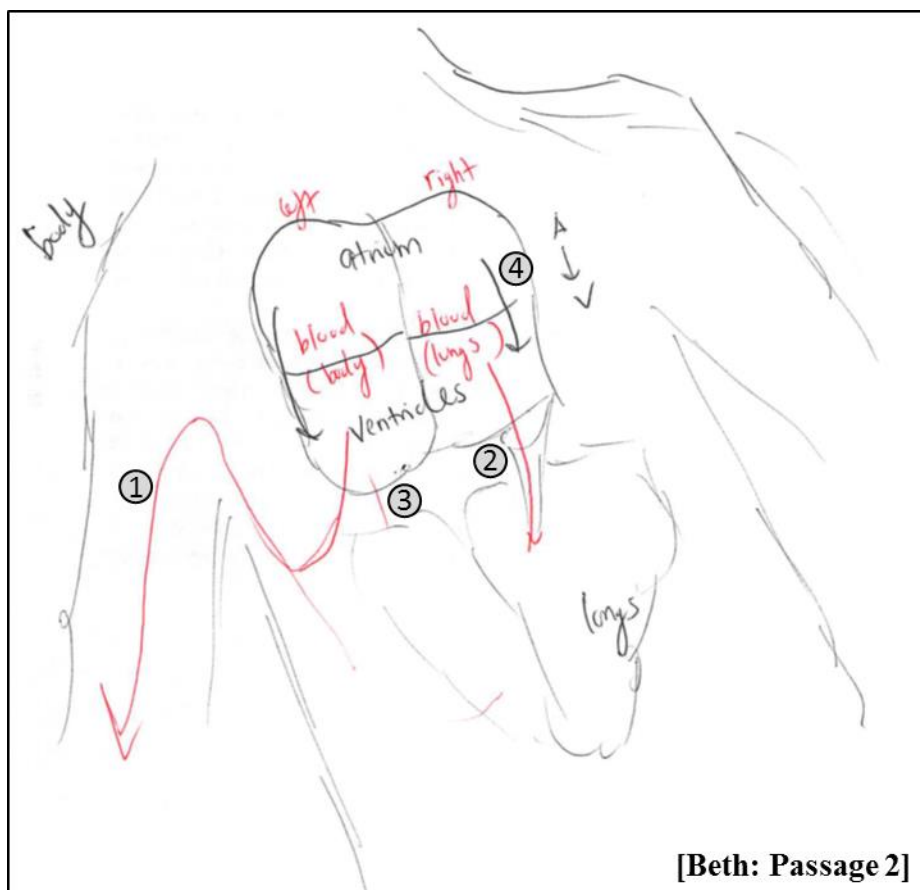
*a. Unfocused content*

It was sometimes difficult to determine if some of Beth's drawings were meant to have a content focus because they included information with no discernable emphasis. Her drawings generally included most or all of the information described in the associated passage but were often arranged in a manner that did not highlight the main ideas of the passage the way that Adam's drawings did. Beth often included an outline of a body in her drawings and always drew the heart inside of it with the lungs underneath. While this may provide the heart and lungs with a physical context in the drawing, this added detail sometimes makes Beth's drawings more complicated and confusing rather than more meaningful (although maybe not from Beth's personal perspective).

Figure 41, below, is Beth's drawing of Passage 2. Her drawing includes the relative positions and names of the heart chambers, and the fact that the left side of the heart pumps blood to the body while the right side pumps blood to the lungs. While complete in the sense that her drawing accounts for all of the information described in the passage, Beth's drawing has some confusing parts to it. First, the red arrow representing the left side of the heart pumping blood to the body is only directed down the left arm (item 1). Second, the red arrow representing blood leaving the right side of the heart is only drawn to the right lung (item 2). Third, there is a red line drawn between the left ventricle and the left lung (item 3), but it is unclear if this line was drawn intentionally or not. Fourth, the arrows drawn, presumably, to represent blood flow from the atria to ventricles are black rather than red like the arrows coming out of the ventricles (item 4).

It is unlikely that Beth had conscious intentions for the irregularity of all of these representations in her drawing. That is, she did not actually believe that the left ventricle only pumps blood down the left arm. However, the lack of systematicity in this drawing, in combination with the arrangement of the organs within the confines of the body outline, make Beth's drawing of Passage 2 much less clear than Adam's (Figure 32). While the same general information can be extracted from Beth's drawing as from Adam's, there is less of a focus on the structure of the heart and where blood is pumped from each side of the heart.

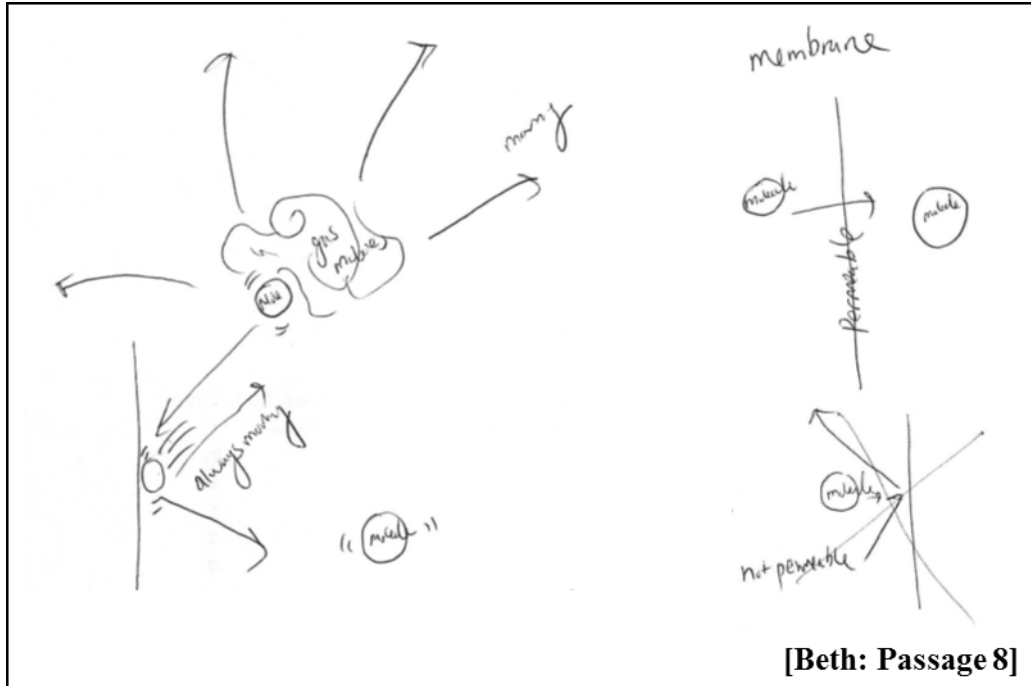




**Figure 41:** Beth's drawing of Passage 2, the general structure of the heart.

Another contrasting case of content focus between Adam and Beth's drawings was with their drawings of Passage 8 about the process of diffusion through a membrane. Remember that Adam contextualized his drawing for this passage by using the walls of a blood vessel to represent the permeable membrane that was described by the passage (Figure 34). Beth, on the other hand, drew multiple representations: a molecule (drawn as a circle) moving and bouncing, another passing through a membrane (a solid line), and another not passing through a membrane (Figure 42). The drawing is reminiscent of a textbook figure of Newton's first law of motion, with no reference to the CVS.

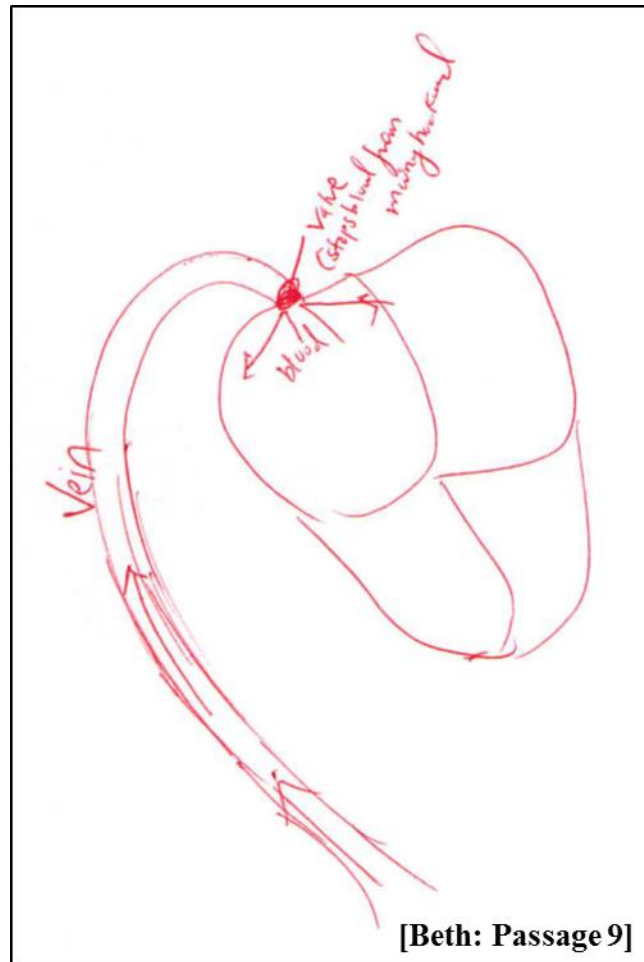
Beth's drawing was not explicitly contextualized within the CVS but, again, it included all of the information described in the passage. To her, these circles could very well represent oxygen or nutrients and the line could represent the wall of a capillary. However, Adam's drawing of two membranes opposite each other with molecules moving between them is much more evocative of the diffusion of nutrients through a capillary wall (Figure 34).



**Figure 42:** Beth's drawing of Passage 8, the process of diffusion through a membrane.

*b. Expand the scope of the text*

Since Beth's drawings tended not to focus explicitly on the main ideas of the passages they represented, they tended to be drawn with a larger scope and include a few other details about the CVS that were not the main foci of the passage. For example, Beth's drawing of Passage 9 that described the function of valves includes an image of the four chambers of the heart (Figure 43). First, there are no valves that separate veins from atria on either side of the heart. There are valves that separate the atria from the ventricles and the ventricles from the arteries, but Beth did not include those in her drawing. Second, if those valves are not shown in the drawing, the heart having four chambers is unrelated to the passage.



**Figure 43:** Beth's drawing of Passage 9, the function of valves.

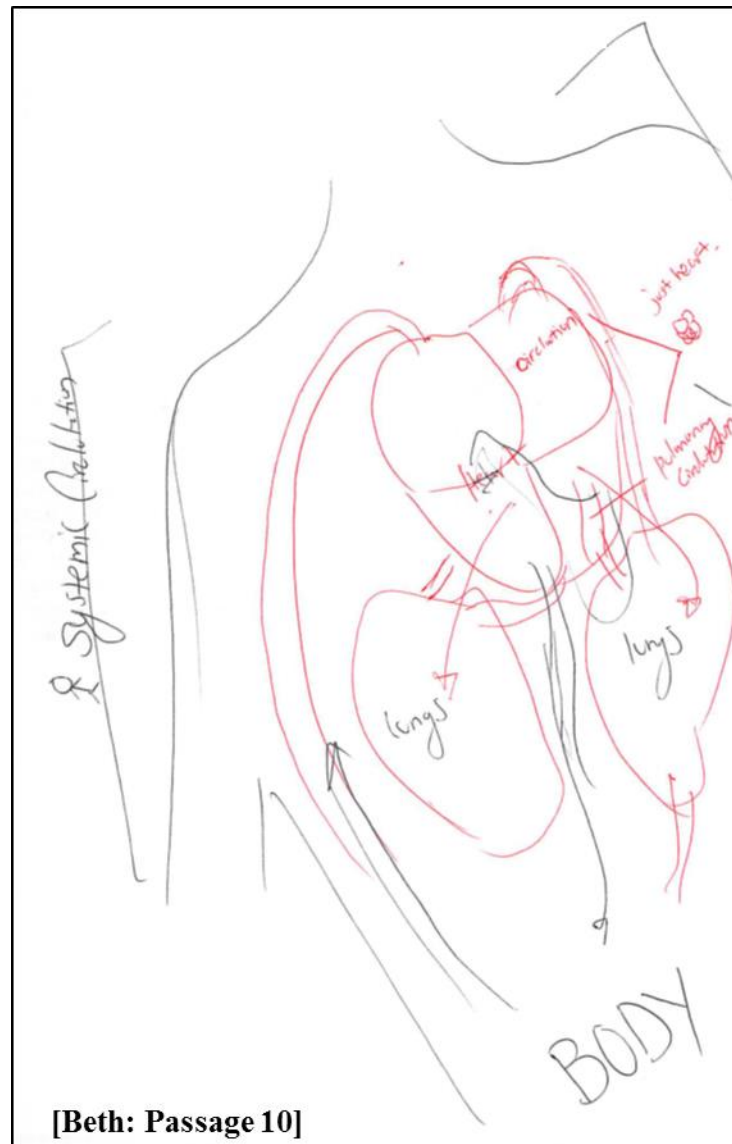
Beth's drawings tended to have a slightly larger scope than the content presented in the CVS passages, likely as a partial result of her drawings not having a strong focus on the passages' main ideas.

*c. Reflect realism*

Beth's drawing of Passage 2, describing the general structure of the heart, was presented earlier and showed that she drew images of the heart and lungs within an outline of a body (Figure 41). Compared to Adam's more abstract drawing, her representation was more realistic. However, this design seemed to have made it more difficult for Beth to represent where the two sides of the heart pumped blood to, seeing how she drew an arrow out of the left ventricle and down the left arm to represent blood being pumped to the body. In addition to other details, this made Beth's drawing not as easily interpretable as Adam's more abstract version.

A comparison of Adam and Beth's drawings of Passage 10 also reveals a difference in the levels of abstraction they used in their drawings. This passage describes the two primary subsystems of circulation. Recall that Adam's drawing was highly simplified with only images of the heart, lungs, and arrows, making the components associated with pulmonary and systemic circulation very distinct. In Beth's drawing, on the other hand, she included details that made her drawing look more realistic but also more cluttered (Figure 44). It is unclear where each of the subsystems begin and end, and where some of the blood vessels are coming from or going to.

Some of the confusion comes from the fact that Beth incorrectly drew blood coming from each ventricle to either the left or the right lung. Furthermore, the use of two lines to outline blood vessels, the separation of the heart chambers, and multiple blood vessels being connected to each lung make Beth's drawing look more realistic and simultaneously unclear.



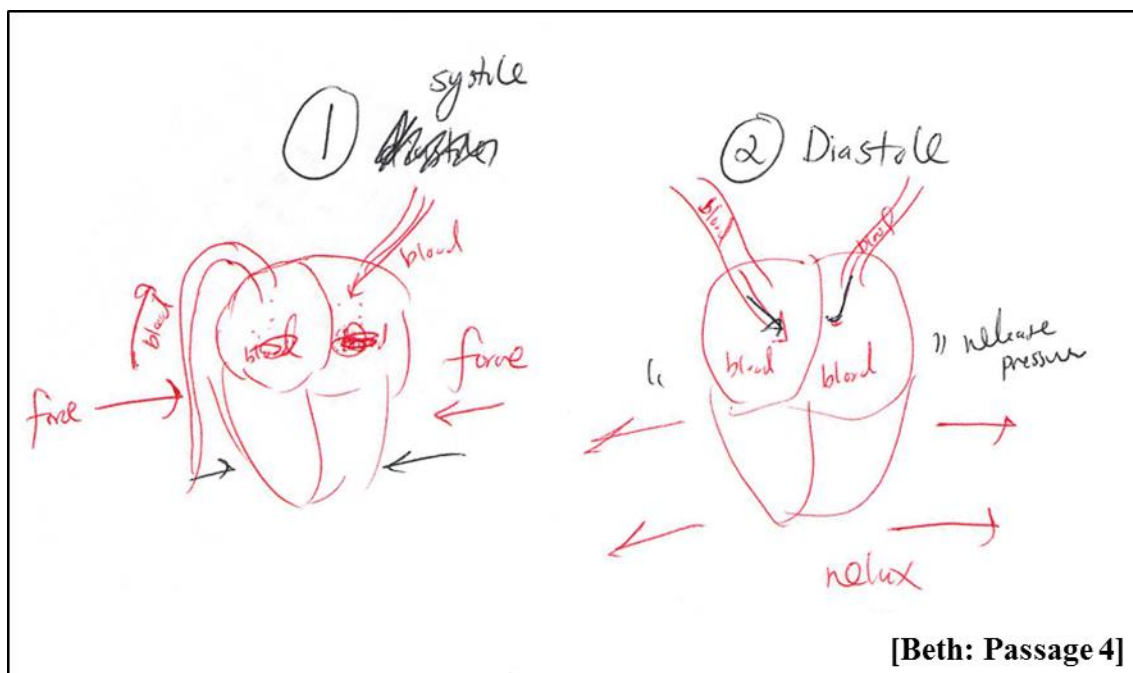
**Figure 44:** Beth's drawing of Passage 10, the two primary subsystems of circulation.

*d. Manage motion and time*

Many of the previous descriptions have shown how several of Beth's drawings were confusing and cluttered by the details and the arrangement of images, but not all of them were uninterpretable. Her drawing of Passage 4, which describes the two phases of muscle contraction, was a clear representation of these phases (Figure 45). She made the decision to separate the two phases with two images of the heart. For systole, she shows the walls of the ventricles constricting, using arrows pointing inward on each ventricle with labels for "force." For diastole, she shows arrows moving away from the ventricles to show the walls of the ventricles relaxing with the labels "relax" and "release pressure." This drawing clearly shows

that systole is the phase where ventricles constrict and diastole is the phase where they relax. What is missing from the drawing of both phases are arteries. As a result, blood is not shown leaving the ventricles in systole. Blood is also not drawn entering the ventricles in diastole.

While Beth's drawing of Passage 4 is missing some details, it is a clear representation of the phases of heart contraction. She did not include images of the lungs or an outline of the body to provide a context for the heart as she did in some of her other drawings. This drawing is concentrated on the phases of heart contraction.



**Figure 45:** Beth's drawing of Passage 4, phases of muscle contraction.

#### e. Summary of Beth's drawings

Beth's drawings tended to represent most of the information presented in their associated passages, although sometimes inaccurately. While all of the information from the passages could potentially be interpreted from the drawings, they were often not the focal points of the drawings. She tended to include an image of the whole body in her drawings, providing a realistic context for the arrangement of structures within the body but also making the drawings cluttered with images. While several of her drawings were inaccurate and difficult to interpret, Beth was occasionally successful in generating clear drawings for certain passages, including the passage that required a representation of a motion that was not just blood flowing.

## 6.5 Discussion

Analyzing self-explanations and drawings from two participants with different learning gains revealed several and a variety of differences in their utilizations of the study strategies. While Adam did not generate a numerous amount of self-explanations, the ones that did generate were consistently meaningful and high-quality, in terms of being accurate, consistent, simple, and precise. They were also meaningful in terms of filling gaps in the text, proposing specific purposes for parts of the system, and fitting the content into the larger scope of the whole CVS. His drawings highlighted the main ideas in the passages and focused only on information that

was relevant to those ideas. His use of abstraction and his ability to clearly represent motion in his drawings also helped to make them easily interpretable.

While Beth generated many more self-explanations than Adam, many of them were lower-quality. She often made explicit efforts to memorize the information by making superficial connections to non-CVS material. Furthermore, when she was not self-explaining to memorize, she sometimes made inaccurate or inconsistent statements that lacked the detail necessary to make the self-explanation meaningful. Her drawings were similarly inaccurate at times. They tended to expand the scope of the information presented in text and be more realistic, but these also made her drawings appear more cluttered and disorganized. Beth was, however, able to clearly represent motion in her drawings, even when tasked with drawing the phases of muscle contraction.

With these case studies we found that even though a student may generate numerous self-explanations, even goal-oriented ones, many of them may not be meaningful or useful for learning. Distinguishing self-explanations from non-self-explanations, and further identifying self-explanations that were goal-oriented, elaborative, or noticed coherence only begins to scratch the surface of understanding different qualities of self-explanation. Similarly, while coding for the number of labels and images in a drawing qualifies the amount of passage-specific detail included, it fails to address other ways in which a drawing can be meaningful for learning.

In this section, we will discuss what factors may influence how learners use the self-explanation and drawing strategies, and what different attributes of their products may have an impact on learning. We conclude with some remarks about other considerations.

#### 6.5.1 Personal backgrounds

---

Students will approach the use of self-explanation and drawing differently, based on their personal experiences. The extent to which a learner has experience with speaking (English), explaining, drawing, visualizing, etcetera, will affect which aspects of self-explaining and drawing they will feel more comfortable with, as well as what they may focus more attention on. For example, if English was not Adam's primary language, then it would be understandable how speaking aloud could be detracting in comparison to drawing because speaking may require more effort from him than writing or drawing. Furthermore, if we consider that Adam was an electrical engineering and computer science major with senior standing, there is a chance that he had more experience with visualizations and designs from his coursework than Beth, who was a sophomore business major.

There are a large number of other factors related to a student's background that can influence how they would use the self-explanation and drawing strategies to learn. Participants in this study varied in gender, age, ethnicity, major, and learning styles. The training sessions conducted at the beginning of each study session were meant to teach each participant how to use each strategy effectively, hopefully reducing the impact of some of these variables. However, participants were still inevitably going to use the study strategies differently.

#### 6.5.2 Attention to aesthetics

---

Another way we tried to reduce differences in study strategy use was to tell participants twice during their training sessions to not be concerned with how their self-explanations sounded or how their drawings looked. They were also encouraged to use the study strategies for themselves and not for the researcher or anyone they might imagine. However, some participants are likely to have still felt inclined to focus on the aesthetics of their self-explanations and

drawings, and less focused on using the strategies to their fullest ability for learning. These potential distractions were expected to impact how effectively participants used the study strategies.

At the same time, considering the fact that Adam mentioned being worried about “explaining or the actual speaking part” of self-explaining and that his self-explanations were high-quality, suggests that perhaps an emphasis on aesthetics could indirectly encourage students to generate more meaningful self-explanations. He may have had a similar attitude towards his drawings. In the same respect, Beth did not seem concerned with how her self-explanations sounded or how her drawings looked, and sometimes generated products that were inaccurate or inconsistent. Adam may have been more careful about the words he chose or the lines he drew, ensuring that the products he generated sounded or looked “good,” while also resulting in higher quality self-explanations and drawings.

### 6.5.3 Self-explaining

---

Regardless of whether or not Adam or Beth attended to the aesthetics of their self-explanations, they each recognized that self-explaining promoted the making of connections. Adam specifically mentioned making connections to prior knowledge, and Beth vaguely mentioned making connections between different concepts. A majority of the time, these “connections” were in fact inferences where the student provided additional information that was not presented by the text.

Now, while all inferences qualify as self-explanations some inferences can be more useful for learning than others. For example, one can introduce something that is only superficially related to the topic, such as some of the associations Beth made strictly for memorization. On the other hand, one can add information from prior knowledge or through reasoning to fill a gap in the text. The first kind of inference might help a learner remember the names of components, but the second is much more likely to help a learner understand why and how parts of the CVS work. Here, we will consider why some attributes of self-explanations might make them useful for learning.

#### *a. Memorizing versus understanding*

One clear distinction between Adam and Beth’s self-explanations were Beth’s explicit attempts to memorize. She tended to introduce new information and make associations while self-explaining passages to help her remember specific terms. These types of utterances, which often connected the CVS material to non-CVS material, were necessarily coded as self-explanations because they introduced information beyond what was presented in the text. This partially accounts for Beth’s high total number of self-explanations, as she made several of these connections.

Compared to some of the high-quality self-explanations that Adam generated, these self-explanations focused on memorizing are not expected to be as useful for learning about the whole CVS. They are likely to help students recall specific terms and even functions (e.g., Beth’s reference to dilating pupils; B.SE2) and can, therefore, be somewhat useful for learning about parts of the CVS. However, memorizing in this manner focuses on isolated pieces of information. Given that the CVS is made up of an intricate network of organs and blood vessels, physically or functionally linked together, a full understanding of the system would be difficult to achieve with isolated pieces of knowledge.

*b. Self-explanations with purpose*

Not all of Beth's self-explanations were focused on memorization, but even these did not give the same impression of high-quality as most of Adam's. I think Beth consciously attempted to make connections with the material as she self-explained. She frequently made inferences, introducing new information in her self-explanations. However, her approach to self-explaining seemed to be more about making connections with any information, rather than CVS-specific information. Connections with non-CVS information would help her to memorize, as discussed above. The remaining connections with CVS-specific information may have helped her to piece together some parts of the system.

Adam's self-explanations were generally high-quality because they were accurate, consistent, simple, and precise, but they were also *specifically* high-quality because they each usually served a purpose for understanding the CVS. The end result of these self-explanations was linking information about the CVS together, and making these types of connections was necessary for learners to develop coherent mental models of the whole system.

Adam sometimes used self-explanations to fill information gaps in the text. Given that text has certain limitations – an inability to describe all aspects and details of the CVS in 11 passages being one of them – being able to fill the gaps is sometimes crucial to comprehending text. By definition, a gap is a break in continuity. With the CVS, specifically, filling a gap inherently links at least two pieces together, on the way to connecting all of the pieces.

Other times, Adam used self-explanations to propose specific purposes for structures, pathways or functions described in the text. That is to say, he explained why something would be structured a certain way, why blood would move from one location to another, or why something would behave in a particular manner. Specifically in the CVS, where every component has survived or developed through evolution to form the highly efficient system, all parts components play essential roles. Therefore, in order to form a full understanding of the CVS, one must understand the purpose of each component. And since these purposes were not always explicitly described in the text, self-explanations that reasoned (or at least made inferences) about them were vital for understanding the CVS.

The last purpose identified that Adam used his self-explanations for was to situate the specific content in passages within the larger CVS. Since the various passages each focus on different aspects of the CVS, it is up to the learners to understand how the passages are related to each other and to the system as a whole. Making these connections out loud in a self-explanation is likely to enforce these connections.

Beth's self-explanations did not emit the same sense of purpose as Adam's. They did not seem to fill many gaps in the text, as they were more likely to add peripheral details or make minor associations. Overall, they usually lacked too much detail and specifics to be purposeful. Of course, Beth did generate a few self-explanations that were insightful and would be helpful for developing an understanding of the CVS, but her typical self-explanation was not as meaningful as many of Adams, and in the end she made several incorrect self-explanations.

*c. Incorrect self-explanations*

While self-explaining passages 2 through 11, in total, Adam generated one self-explanation with incorrect information in it. Beth generated 8. These statements were necessarily



coded as self-explanations because they included information not presented in the text. Across all participants' self-explanations for passages 2 and 3, the average was 0.7 incorrect self-explanations per participant. This small number was unlikely to have an effect on the correlations conducted in the last chapter. However, across all participants and all passages, the effect of incorrect self-explanations should be accounted for separately.

As seen with Beth's self-explanations of the final passage in the booklet, and in her comments about the challenges of self-explaining, this study strategy does not help to clarify misunderstandings. Beth stated that she was confused about the material and then further confused herself as she self-explained. Her inaccurate responses on the post and delayed assessments indicate that these incorrect interpretations of the text remained in her mental model when she finished her study session.

Considering Beth's case, self-explaining incorrect information may seem like it would only have negative effects on learning. Obviously, one of the worst case scenarios is when a learner generates an incorrect self-explanation because he or she misinterpreted the text and then comes to believe this misinterpretation to be true. However, consider the following points: (a) people sometimes accidentally say one word when they mean and are thinking another, (b) stating an incorrect self-explanation aloud may lead a person to recognize the fault in their statement and revise it (Chi, 2000), (c) some parts of the incorrect self-explanation may be correct and meaningful. Identifying these instances is beyond the scope of our data, but it was for these reasons that incorrect self-explanations were not excluded from the analyses.

#### 6.5.4 Drawing

---

Both Adam and Beth expressed preferences for drawing over self-explaining at the end of their study sessions. Adam was not specific about how drawing was better than self-explaining, other than it being a visual strategy for learning a topic that was visual in nature. However, Beth mentioned that she could physically refer back to her drawing while studying the associated passage (but could not refer back to prior self-explanations).

In this section, we will discuss why some of the attributes identified from Adam and Beth's drawings might make them useful for learning.

##### *a. Content focus*

Adam's drawings tended to concentrate on the main ideas presented in their associated passages. Not only were details in his drawing minimal but the drawings were designed in such a way that the main ideas in the passage seemed to be brought to the foreground. On the other hand, only some of Beth's drawings concentrated on the main ideas of the passages. She sometimes included peripheral details in her drawings that made them more difficult to interpret. Although, perhaps to Beth these details were not distracting and the main ideas of the passage stood out to her.

In general, just because a participant's drawing does not appear to focus on any particular concept, does not mean that the participant did not focus on it while he or she was studying. Remember, all students were asked to self-explain and draw for themselves. However, drawings that were clear and focused to an observer, like Adam's drawings, were also likely to be easier to review by the learner who drew them. Especially for more complex passages with a lot of information, seeing a clear compilation of all of the pieces together can help a learner see how they fit together. Thus, focusing the content of drawings on the main ideas of each passage may help the learner to actively highlight the important concepts in the passage for him or herself.

Ultimately, a key benefit to drawing is being able to refer back to one's finished product. Thus, it is important to have a product that is easily interpretable by the learner so that he or she can review the main ideas from the passage before moving on to the next. Perhaps, Adam proactively generated his drawings in a clear manner with the intent of reviewing his complete drawing when he was finished, whereas Beth generated her drawings reactively. This is reminiscent of the expert students in Martin & Schwartz (2009), whose drawings were better planned-out for future use than the novice students', and the novice students in Kindfield (1994), whose drawings tended to include extraneous details.

*b. Scope*

While the expansion of the scope of a passage seems to be productive when self-explaining, the opposite may be true for drawings. In some cases, including additional content in one's drawing can add more meaning as it does with self-explanations. However, there are also cases where the additional information makes other parts of the drawing less clear and actually reduces meaning. For example, a subsequent image added to a drawing can obstruct or cross the boundaries of a previous image. This is not the case with self-explanations because a subsequent self-explanation will not change the meaning of a previous one (unless it was intentionally meant to refer back to an utterance and do so). Therefore, while expanding the scope of the information presented in the text can increase meaningfulness in a self-explanation, it might not in a drawing. Adam's tendency to restrict the scope of his drawings to the content presented in the passage seemed to make his drawings more focused and clear, and clarity is important for reviewing concepts and seeing the accumulation of all of the pieces of information put together.

*c. Level of abstraction*

In the context of this study, being able to abstract and generalize information from a passage into a drawing may be a sign of deeper understanding. Adam had to understand which features were extraneous and could be excluded from his drawing, and, likewise, which features were necessary for portraying the main ideas of the passage. However, just because a student did *not* create a more abstract version of their drawing, does not mean that they would not be able to if asked. Abstraction can be more of a design choice rather than a reflection of one's understanding of the CVS. This will be discussed in more detail in the Meta-representational competence section below.

*d. Motion and time*

Motion and time are relatively more difficult to represent in a static drawing than a structure or pathway (and motion and time are often associated with a function). Our two case study participants were successful at representing blood flow, the diffusion of molecules, and the phases of muscle contraction. However, those were the only aspects of the CVS text that involved motion. Compared to the representations that students generated of an object slowing, stopping, and then speeding up in diSessa et al. (1991), the representations of the types of motion described in this CVS text were not as complex. There was also neither motivation nor need for students to revise, perfect, or abstract their representations of these motions, as there were many other aspects to learn from the text.

Nevertheless, there were interesting differences found between Adam and Beth's drawings of the phases of muscle contraction. For example, Adam used arrows pointing inward

to represent the ventricles contracting starting from the ventricle walls and moving towards the septum (Figure 38). Meanwhile, Beth also used arrows pointing inward, but her arrows started from outside of the ventricles and pointed towards the outside walls of the ventricles (Figure 45). From an observer's perspective, Adam's drawing looks like the walls of the ventricles are moving towards the septum, while Beth's drawing looks like the walls are being *pushed* towards the septum by an outside force. These factors are also related to meta-representational competence, discussed next.

*e. Meta-representational competence*

The level of abstraction a learner uses in a drawing and how the learner chooses to represent motion may be more of an indication of meta-representational competence than understanding of the CVS. Imagine student A who drew a well-organized and clear drawing and student B whose drawing contains all important pieces of information but is not as clear. In one situation, student A may have a better understanding of the pieces of information and how they fit together and can, therefore, better organize the information into a clear drawing. Student B, on the other hand, might understand some pieces of the information (enough to include them individually in a drawing) but does not have a fully coherent understanding of the system and has difficulty arranging the information. In another situation, student A and student B both understand all of the information equally, but student B does not know how to translate the information into a drawing. In a third situation, student A is the type of learner who likes to organize information and student B is not.

Had Adam drawn Passage 2 with the heart and lungs inside of an outline of a body (instead of the body to the left and the lungs to the right; Figure 32), he still could have shown that the left side of the heart pumps to the body and the right side pumps to the lungs, but would have had to represent it in a different way (e.g., with arrows pointing all around the body and/or to a label indicating the rest of the body). His more abstract arrangement simplified and highlighted the fact that blood flows from the left side of the heart to the body and from the right side of the heart to the lungs. But it is unknown if Adam intentionally generated his drawing for these effects. Others with less experience or skill in representing information may default to an arrangement that is more realistic, and automatically place the heart and lungs within the outline of a body.

#### 6.5.5 Other comments

---

*a. Preconceptions*

Many of the claims presented in this chapter were loosely based on two premises: 1) that Adam learned more about the CVS than Beth and 2) that this was because he used the self-explanation and drawing study strategies more effectively than she did. While, in general, I do believe both of these to be true, I believe that each student's understanding of the CVS upon entering the study were significantly different. They were different in a complicated manner that would not be indicated by pre-test scores.

Recall that Adam responded to 11 of the 18 questions on the pre-test with "I don't know," and was likely to have taken AP Biology at least a year earlier than Beth. Beth only responded with "I don't know" once. However, their pre-test scores were very similar to each

other. This means that most of the questions that Adam responded to with “I don’t know,” Beth responded to incorrectly.

What I am proposing is that Beth having (incorrect) answers to many pre-test questions may be an indication that she had an inaccurate mental model of the CVS coming into this study, while Adam may have simply had an incomplete model. Thus, it would have been more difficult for Beth to correct the inaccurate parts of her model than for Adam to fill the missing pieces of his (Chi, 2000). This idea relates to the discussion of violations introduced in Part 1 of this dissertation. The inaccurate parts of Beth’s preconceived model might be embedded within a network of other information that somehow made her model work. “Fixing” this part would affect all knowledge connected to it. Therefore, even if Beth used the self-explanation and drawing strategies as well as Adam did, it would have been more difficult for her to achieve a correct model of the CVS than for Adam to.

*b. Self-explain and draw to learn or vice versa?*

Beth mentioned in her opinions about self-explaining that it was difficult for her to generate self-explanations if she did not understand the material first, otherwise her self-explanations would further confuse her. While this statement sounds contradictory to the principle of self-explaining to learn, it may have been a feeling many other participants shared.

Self-explanation is meant to encourage learners to form connections either between the text and prior knowledge or between different parts of the text. However, if the learner does not understand a piece of the text, he or she cannot form a meaningful connection to it. The learner might form a connection to the material, but it is more than likely to be a superficial one, such as one of the low-quality self-explanations Beth generated during her study session. The same is true of drawing. If a learner does not understand a piece of the text, how can he or she draw it? Thus, while self-explaining and drawing are intended to encourage learners to generate connections, a threshold of understanding – from reading the text alone – is necessary for a meaningful connection to be made.

## 7 OVERALL DISCUSSION

---

In this section, we consider the findings from this study altogether to address the research questions posed at the beginning of this dissertation. First we will review the findings from each part of this study, then address each question individually. We will conclude with some discussion of the limitations of this study and some concluding thoughts.

### 7.1 Summary of Findings

---

The first unexpected finding from our study was that participants in the Rewrite group appeared to learn more about the CVS than participants in SE group on many formal assessment measures. This is attributed to the fact that these participants likely spent a substantial amount of time with the CVS text studying, in some manner, in silence.

#### 7.1.1 Part 1: Formal Assessments

---

Overall pre-test scores showed that the average student in each experimental group entered this study with approximately the same level of understanding of the CVS. After using their assigned study strategies to study the CVS text, however, participants in the self-explanation group seemed not to learn as much as participants in the other groups, including the rewrite group. Their average drawing mental model scores did not improve as much as the other groups, especially the Draw group.

While most of the differences in free-response scores were not statistically significant, the same pattern was seen in scores on the different question sets and on the post and delayed tests. The self-explanation group tended to fall behind the other groups mostly in learning about structures and pathways. Finally, the average participant in each group labeled the same number of items correctly on the figure interpretation task, but participants in the Draw group tended to draw more incorrect arrows to indicate the direction of blood flow on the figure.

#### 7.1.2 Part 2: Use of Study Strategies

---

Different types of self-explanation and drawing formats were found to have different relationships with SPF sub-scores. In general, the number of goal-driven self-explanations a learner generated was correlated with each of their SPF sub-scores, but strongly with functions sub-scores in both the SE group and the Combined group. Furthermore, mentioning functions in a self-explanation was strongly correlated to functions sub-scores in each group, but also with pathways sub-scores in the SE group and with structures sub-scores in the Combined group.

In regards to the analyses of student drawings, labels were found to be negatively correlated with SPF sub-scores in the Draw group, but positively correlated with sub-scores in the Combined group. In addition, Images were strongly correlated with structures sub-scores in both groups, and also with functions sub-scores in the Combined group.

#### 7.1.3 Part 3: Case Studies

---

Closer analyses of two participants in the Combined group showed that the participant who appeared to learn more about the CVS, tended to generate self-explanations that filled key gaps in the text, propose specific purposes for components mentioned in the text, and expanded the scope of the content in the passage. Moreover, his drawings highlighted the main ideas of the passage, focused only on relevant information, and displayed some abstraction. In comparison,

the participant who learned an average amount about the CVS (based on relative post and delayed test scores) generated many self-explanations that were focused on memorizing information. She tended to be vague in her inferences, present peripheral details, and occasionally make incorrect inferences. Her drawings were often unfocused, also presenting peripheral details, and tended not to stray from realism.

## 7.2 A Return to the Research Questions

---

### 7.2.1 Research Question #1

---

*As independent study strategies for overall learning about the cardiovascular system from text, how effective is drawing in comparison to self-explaining?*

As hypothesized, students who drew while studying the CVS text tended to learn more than students who self-explained. At the beginning of the study, the average participant across all groups drew a model of the CVS that contained only one loop that did not include the lungs. After studying, the average participant in each group drew two loops (including the lungs) except in the SE group, who only drew a single loop with lungs. Thus, all participants demonstrated noticeable improvement in their drawn models after using their assigned study strategies to learn from the text, but those in the SE group did not improve as much as others. Furthermore, although a drawing cannot fully represent a student's internal mental model, I believe that these drawings are, to some degree, an extension of them and show that many students in the SE group did not understand the two-loop system after studying the material.

### 7.2.2 Research Question #2

---

*Do students who draw learn about different aspects about the cardiovascular system than students who self-explain? For example, do they learn better about more visual aspects of the system?*

Students who had the opportunity to draw while studying tended to understand more about structures and pathways than students who self-explained, but not about functions. Obviously, it is easier to show structures and pathways in images than describe them in text or verbally in words (Larkin & Simon, 1987), but it was uncertain if this act of drawing and seeing these images for oneself would truly impact learning. Students who self-explained while studying could easily internally visualize these structures and pathways to make up for the fact that they could not draw. However, given that the CVS is complex and the components involved are numerous, it is likely that drawing allowed learners to offload some of the information they had to hold in their working memories, freeing some capacity to better process the information (Scaife & Rogers, 1996).

Overall, it was anticipated that students who drew during their study sessions would perform better on visual assessment tasks than students who did not, because learning in a visual mode was expected to help with transfer of learning to assessment tasks formatted in the same mode (Barnett & Ceci, 2002). If we consider the drawing task, this may have been the case, but I also believe that students in the Draw group out-performed those in the SE group because they had better internal mental models of the system, which had little to do with the matching modes supporting transfer of learning. Granted, students who drew while studying may have found clearer ways of representing the information in the system by the time they needed to draw on the post-test and may have formed subconscious connections with and between information

through the act of drawing. But, remember, the difference between the average student in the Draw group and in the SE group was one loop versus two. This distinction is usually easily distinguishable in a drawing (even in a messy one), and the benefits of consistency between mode of study and mode of testing are unlikely to create a difference like one versus two loops in the system.

Finally, the figure interpretation task, while visual in nature, was a different type of visual task. Instead of generating an inscription, learners were asked to interpret one. Students in the Draw group did not perform better than those in the SE group. In fact, they performed worse when asked to draw the direction that blood flowed through the system in the figure. I proposed that the reason for this was that, during their study sessions, students in the Draw group almost always represented blood coming out of ventricles in a downward direction, and did so repeatedly for different passages. However, in the anatomically correct figure for the task, blood would be pumped out of the ventricles in an upward direction. Thus, students who had familiarized themselves with their drawn version of the direction of blood flow, would not recognize this or would not know how to resolve the conflict.

In summary, drawing was better than self-explaining from text for learning about structures and pathways but not functions in the CVS. As a result, drawing was likely better for developing a two-loop mental model of the system. Finally, practice drawing may have helped learners to translate that model into an inscription. A shortcoming of this study strategy may be that learners can become attached to their version of the inscription and have trouble reconciling it with other versions.

### 7.2.3 Research Question #3

---

*Does self-explaining and drawing together have additive or synergistic effects?*

No. On some measures, drawing alone was better than the combination of self-explaining and drawing while studying the CVS text. Using both study strategies was originally hypothesized to be better than using only one of them because learners could gain from the combined benefits of each. However, students in the Combined group tended to take the longest to complete the CVS text booklet, and may have experienced the most fatigue. In addition, on average, they spent less time self-explaining than the SE group, and less time drawing than the Draw group. It may be that they felt, at times, as though using both strategies was somewhat repetitive. For example, they might not be motivated to generate a drawing that was very detailed after spending a substantial amount of time self-explaining the same material.

Participants in the Combined group generally had functions sub-scores that were a bit higher than the other groups, and structure and pathway sub-scores that were higher than the SE group and sometimes higher than the Draw group. As proposed in Part 1, perhaps the participants in the Combined group benefitted from the focus that both self-explaining and drawing placed on functions and the focus that drawing placed on structures and pathways, but was also detracted from structures and pathways by self-explaining. Thus, self-explaining and drawing together was sometimes better than only doing one or other when studying aspects that each were shown to be effective for learning. Other times, using both strategies may have shifted attention away from certain categories if they were not supported by each strategy separately.

#### 7.2.4 Research Question #4

---

*What types of activities do students engage in when they are drawing to learn and do any of those activities correlate to learning gains? In other words, while “Good students” tend to self-explain more when they are asked to think aloud (Chi et al., 1989), do they tend to do something analogous in their drawings?*

The number of passage-specific images that students included in their drawings was strongly correlated with post and delayed test scores. The strength of these correlations were comparable to those found between self-explanations and the same scores. Therefore, a similarity exists between images and self-explanations where the more of each that a student generates while studying, the more they tend to learn.

As seen with the case studies presented in Part 3, it is important that the images a learner includes in his or her drawing be specific to the content presented in the passage that the drawing is meant to represent. Perhaps in the same manner that some of Beth’s self-explanations were less meaningful than others, some images can also be less meaningful than others for learning. I presume that these passage-specific images embody the fact that the student read the information presented in the text, considered how to represent it as an image, and considered how the image would fit in with the rest of the drawing.

#### 7.2.5 Research Question #5

---

*What specific attributes of self-explanations and drawings make them meaningful for learning?*

In general, self-explanations that are accurate, consistent, simple, and precise tend to be more meaningful for learning. Specifically, self-explanations need to serve a purpose for the learner. Participants were trained to use the self-explanation strategy to make connections within the text, make connections to prior knowledge, make predictions, etc., and one might assume that doing any of these things would be sufficient for supporting learning. However, if the student did not see that the purpose of making connections with the text was to help them understand all of the parts of the system, their self-explanations tended to lack the focus or detail needed to make them meaningful. Self-explanations that served a purpose tended to fill gaps of information in the text, infer specific purposes for components, and expand the scope of the text to include more parts of the larger system.

In general, drawings that were clear and focused were likely to be more meaningful for learning. Considering that part of the benefit of using drawing as a study strategy is having a physical record of previous ideas that were considered, being able to refer back to one’s drawing and review the culmination of all of the images together would support learning. Specifically, a drawing that focuses on the main ideas of the passage, without including much peripheral detail, allows a learner to review the important information that they processed from the text.

### 7.3 Limitations

---

This dissertation presents an analysis of a subset of a large dataset collected. While I trust that the sample of data used in this dissertation is representative of the entire set, if every self-explanation and every drawing that the 96 participants in this study generated was included in the analyses, the results may have been somewhat different. Other limitations of this study are described in this section.



There is very little data on the participants in the Rewrite group and what they were doing during their independent study sessions. It was surprising that being asked to copy text seemed to be more effective for learning than self-explaining, a study strategy that has been shown repeatedly in the literature to promote deep learning. Retrospective interviews would have been very informative of what other study techniques these students were using during their silent times with the CVS text.

The formal assessment questions used in this study have been used in prior research and were adequate for identifying some gaps in student understanding. However, a verbal assessment in the form of a semi-structured interview would have provided more reliable accounts of student understanding of the CVS before and after studying the text. On several occasions, student responses to free-response items were unclear, and being able to ask follow-up questions could have changed how their understanding of the system was perceived. Interviews would have been especially informative at the time of the pre-test, as many students may not have been aware of the extent to which they should draw details of the CVS, not knowing how much detail would be included in the CVS text they would see later. Thus, some may have understood the double loop pathways in the beginning, but did not include them in their initial drawings.

Aside from the need for additional data, additional analyses can be conducted with this dataset. More thorough process analyses can be conducted with individual students to try to track when, and perhaps how, they came to understand each part of the CVS and the roles they play in the system. This would be interesting to follow through to the post test and delayed test responses to see if the ways that students use self-explanations and drawings can predict what aspects of the system will be forgotten by or retained until the delayed test. Overall, a more detailed method of analysis is required for understanding the complexities of the contents of student self-explanations and drawings in trying to understand what they mean, what purposes they serve the learner, and how they impact learning.

Finally, much background data has yet to be processed that can be informative of how gender, major, or learning style can influence how a student chooses to use the self-explanation and drawing strategies with the CVS text, and also how effective each are in promoting learning for different groups of learners. While there are some limitations to this study, there is much we can learn from it and even more we can use to build from for further research.

## 8 CONCLUSION

---

The findings from this study clearly show that drawing is an effective study strategy for learning about the CVS, and more effective than self-explaining. Considering the fact that some college courses are now being flipped to get students more actively involved in their own learning, during what would be their normal lecture hours, it only makes sense to consider ways for students to be more actively involved in their learning outside of the classroom as well. Self-explanation is an excellent strategy for actively learning from text in many different contexts, including biology, while drawing is an excellent strategy for actively learning from text in many different contexts, *especially* biology. Biology is filled with visual content that is getting more and more complex as advances in technology allow it to.

Drawing can allow students to visually make physical connections between pieces of information in their inscriptions as they simultaneously form internal connections in their mental models. Decisions about what information to represent and not represent provide learners with opportunities to identify main ideas from pools of information. Furthermore, drawing something interpretable requires learners to coordinate aspects of the drawing and ideas to effectively transmit the information in a manner that is clearly interpretable to oneself while reviewing or to another person while communicating.

Promoting drawing as a learning strategy can prepare biology students for the future representations that they will both generate and encounter in their future careers, and other students for the science representations they will find in the media or in medical pamphlets. For all students, practicing and learning to represent information can develop meta-representational competence so that they can understand the purpose of all types of representations in all fields. Further research into how students can draw to most effectively promote learning can only help to produce these benefits.

## 9 REFERENCES

---

- AAAS. (2011). *Vision and Change in Undergraduate Biology Education: A Call to Action*. Washington, DC.
- Ainsworth, S., & Burcham, S. (2007). The impact of text coherence on learning by self-explanation. *Learning and Instruction, 17*(3), 286–303.
- Ainsworth, S., Galpin, J., & Musgrove, S. (2007). Learning About Dynamic Systems by Drawing For Yourself and For Others. In *EARLI conference, Budapest*.
- Ainsworth, S., & Loizou, A. (2003). The effects of self-explaining when learning with text or diagrams. *Cognitive Science, 27*, 669–681.
- Amin, T., Smith, C., & Wisner, M. (2014). Student conceptions and conceptual change: Three overlapping phases of research. *Handbook of Research in Science Education, 2*, 57–81.
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*(4), 612–637.
- Bielaczyc, K., Pirolli, P. L., & Brown, A. L. (1995). Training in Self-Explanation and Self-Regulation Strategies: Investigating the Effects of Knowledge Acquisition Activities on Problem Solving. *Cognition and Instruction, 13*(2), 221–252.
- Bowen, G. M., & Roth, W. M. (2002). Why students may not learn to interpret scientific inscriptions. *Research in Science Education, 32*(3), 303–327.
- Bowen, G. M., Roth, W. M., & McGinn, M. K. (1999). Interpretations of graphs by university biology students and practicing scientists: Toward a social practice view of scientific representation practices. *Journal of Research in Science Teaching, 36*(9), 1020–1043.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school*. National Academy Press.
- Brewer, W., Chinn, C., & Samarapungavan, A. (1998). Explanation in Scientists and Children. *Minds and Machines*.
- Chan, J., & McDermott, K. (2007). The Testing Effect in Recognition Memory: A Dual Process Account. *Learning, Memory, 33*(2), 431–437.
- Chi, M. (2000). Self-explaining: the Dual Processes of Generating Inference and Repairing Mental Models. In *Advances in instructional psychology* (pp. 161–238).
- Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In *Advances in instructional psychology* (pp. 161–238).

- Chi, M. T. H. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. Routledge.
- Chi, M. T. H. (2009). Active Constructive Interactive: A Conceptual Framework for Differentiating Learning Activities. *Topics in Cognitive Science*, 1(1), 73–105.
- Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13(2), 145–182.
- Chi, M. T. H., Leeuw, N. D., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, 18(3), 439–477.
- Chi, M., & VanLehn, K. (1991). The Content of Physics Self-Explanations. *The Journal of the Learning Sciences*, 1(1), 69–105.
- Chi, M. T., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science*, 25(4), 471–533.
- Cox, R. (1997). Representation interpretation versus representation construction: a controlled study using switchERII. In B. du Boulay & R. Mizoguchi (Eds.), *Artificial Intelligence in Education, 1997: Knowledge and Media in Learning Systems (Proceedings of the 8th World Conference on Artificial Intelligence in Education Society)* (pp. 434–444). Amsterdam: IOS Press.
- Cox, R., & Brna, P. (1995). Supporting the use of external representations in problem solving: The need for flexible learning environments. *Journal of Artificial Intelligence in Education*, 6(2/3), 239–302.
- Cromley, J., Newcombe, N., & Bergey, B. (2011). Effectiveness of Student-Constructed Diagrams and Self-Explanation Instruction. Retrieved from [http://spatiallearning.org/resource-info/showcase/showcase\\_index.html](http://spatiallearning.org/resource-info/showcase/showcase_index.html)
- Dirks, C. (2011). *The current status and future direction of biology education research (National Research Council Commissioned Paper)*.
- diSessa, A. (2002). Students' criteria for representational adequacy. *Symbolizing, Modeling and Tool Use in Mathematics Education*, 105–129.
- diSessa, A. A. (2002). Why “conceptual ecology” is a good idea. In *Reconsidering conceptual change: Issues in theory and practice* (pp. 28–60). Springer Netherlands.
- diSessa, A. A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction*, 22(3), 293–331.

- diSessa, A. A., Hammer, D., Sherin, B., & Kolpakowski, T. (1991). Inventing graphing: Meta-representational expertise in children. *The Journal of Mathematical Behavior*, *10*, 117–160.
- Friedman, S. E., & Forbus, K. D. (2011). Repairing incorrect knowledge with model formulation and metareasoning. In *IJCAI International Joint Conference on Artificial Intelligence* (pp. 887–893).
- Gentner, D., & Stevens, A. L. (1983). *Mental Models*. (D. Gentner & A. L. Stevens, Eds.). Hillsdale, NJ: Erlbaum.
- Gilbert, J. (2005). Visualization: A metacognitive skill in science and science education. *Visualization in Science Education*, 9–27.
- Hatano, G., & Inagaki, K. (1986). *H. W. Stevenson, H. Azuma, & K. Hakuta* {(Eds.)}, *Child development and education in Japan: A series of books in psychology*. New York, {NY}: {W.H.} Freeman.
- Hmelo, C. E., Holton, D. L., & Kolodner, J. L. (2000). Designing to Learn About Complex Systems. *Journal of the Learning Sciences*, *9*(3), 247–298.
- Hmelo-Silver, C. E., & Pfeffer, M. G. (2004). Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. *Cognitive Science*, *28*(1), 127–138.
- Jacobson, M., & Wilensky, U. (2006). Complex Systems in Education: Scientific and Educational Importance and Implications for the Learning Sciences. *The Journal of the Learning Sciences*, *15*(1), 11–34.
- Keil, F. (2006). Explanation and understanding. *Annual Review of Psychology*, *57*, 227.
- Kindfield, A. C. H. (1994). Biology diagrams: Tools to think with. *Journal of the Learning Sciences*, *3*(1), 1–36.
- Kindfield, A. C. H., & Singer-Gabella, M. (2010). Inscriptional Practices in Undergraduate Introductory Science Courses: A Path toward Improving Prospective K-6 Teachers. *Journal of the Scholarship of Teaching and Learning*, *10*(n3), 58–88.
- Klionsky, D. (2001). Constructing Knowledge in the Lecture Hall. *Journal of College Science Teaching*, *31*(4), 246–51.
- Kloser, M. (2010). The Implications of the Unique Nature of Biology for Biology Education. A Paper for the Annual Meeting of NARST.
- Knight, J. K., & Wood, W. B. (2005). Teaching more by lecturing less. *Cell Biology Education*, *4*(4), 298–310.

- Kozma, R. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction, 13*(2), 205–226.
- Kozma, R., & Russell, J. (2005). Students Becoming Chemists: Developing Representational Competence. *Visualization in Science Education*, 121–145.
- Larkin, J. H., & Simon, H. A. (1987). Why a Diagram is (Sometimes) Worth Ten Thousand Words. *Cognitive Science, 11*(1), 65–100.
- Latour, B. (1990). Drawing things together. *Representation in Scientific Practice*, 19–68.
- Linn, M. C., & Eylon, B.-S. (2011). Making Visualizations Valuable. In *Science Learning and Instruction: Taking Advantage of Technology to Promote Knowledge Integration* (p. 186–). New York: Routledge.
- Martin, L., & Schwartz, D. (2009). Prospective adaptation in the use of external representations. *Cognition and Instruction, 27*(4), 370–400.
- Mayr, E. (2004). The Autonomy of Biology. In *What makes biology unique? Considerations on the autonomy of a scientific discipline* (pp. 11–38). Cambridge: Cambridge University Press.
- McNamara, D. S. (2004). SERT: Self-Explanation Reading Training. *Discourse Processes, 38*(1), 1–30.
- O'Donoghue, S. I., Gavin, A. C., Gehlenborg, N., Goodsell, D. S., Hériché, J. K., Nielsen, C. B., ... Shattuck, D. W. (2010). Visualizing biological data—now and in the future. *Nature Methods, 7*, S2–S4.
- Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science*.
- Roth, W. M., Bowen, G. M., & McGinn, M. K. (1999). Differences in graph related practices between high school biology textbooks and scientific ecology journals. *Journal of Research in Science Teaching, 36*(9), 977–1019.
- Roth, W. M., & McGinn, M. K. (1998). Inscriptions: Toward a theory of representing as social practice. *Review of Educational Research, 68*(1), 35.
- Roy, M., & Chi, M. (2005). The self-explanation principle in multimedia learning. In *The Cambridge handbook of multimedia learning* (pp. 271–286).
- Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies, 45*, 185–213.

- Schönborn, K. J., & Anderson, T. R. (2006). The importance of visual literacy in the education of biochemists. *Biochemistry and Molecular Biology Education*, 34(2), 94–102.
- Schönborn, K. J., & Anderson, T. R. (2009). A model of factors determining students' ability to interpret external representations in biochemistry. *International Journal of Science Education*, 31(2), 193–232.
- Steele, C., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69(5), 797.
- Stieff, M. (2011). Improving representational competence using molecular simulations embedded in inquiry activities. *Journal of Research in Science Teaching*, 48(10), 1137–1158.
- Stieff, M., Hegarty, M., & Deslongchamps, G. (2011). Identifying Representational Competence With Multi-Representational Displays. *Cognition and Instruction*.
- Sweller, J., Van Merriënboer, J. J., & Paas, F. G. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, 10(3), 251–296.
- Van Meter, P. (2001). Drawing construction as a strategy for learning from text. *Journal of Educational Psychology*, 93(1), 129.
- 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., & Garner, J. (2005). The Promise and Practice of Learner-Generated Drawing: Literature Review and Synthesis. *Educational Psychology Review*, 17(4), 285–325.
- VanLehn, K., Jones, R., & Chi, M. T. H. (1992). A model of the self-explanation effect. *The Journal of the Learning Sciences*.
- Wong, R. M. ., Lawson, M. J., & Keeves, J. (2002). The effects of self-explanation training on students' problem solving in high-school mathematics. *Learning and Instruction*, 12(2), 233–262.

## 10 APPENDIX

---

### 10.1 Appendix A: Screening Survey

---

#### **IB35AC Biology Educational Research Project (Lam)**

If you consent to the terms below, you may continue to the short survey on the next page.

**\* Required**

#### Consent to Participate in Screening Survey

##### INTRODUCTION AND PURPOSE

My name is Diane Lam, and I am a graduate student in the Graduate Group in Science and Mathematics Education (SESAME) program at UC Berkeley, and I am working with my faculty advisor, Professor Angelica Stacy in the Graduate School of Education and two research assistants, Kristina Yim and Michael DeChenne. We would like to invite you to take part in our research study which investigates strategies that students can use to help them learn from reading text about biology.

##### PROCEDURES

If you agree to participate in our research, we will ask you to complete the attached online survey. The survey will involve questions about learning styles, past experiences with science and biology, and basic demographic information. It should take less than 10 minutes to complete. After you submit your responses to the screening survey, we will determine if we can or cannot invite you to participate in the study, and will notify you within 2 weeks via the email address that you provide at the end of the screening survey.

##### BENEFITS

There is no direct benefit to you from taking part in this screening survey, but we hope that results from the main study will inform students and educators of how learners may best learn from texts about science topics.

##### RISKS/DISCOMFORTS

As with all research, there is a chance that confidentiality could be compromised; however, we are taking precautions to minimize this risk.

##### CONFIDENTIALITY

Your responses to the screening survey will be handled as confidentially as possible. If you do not end up participating in the main part of the study, your survey data will be destroyed once participants have been chosen. If you are chosen to participate in the main study, your screening survey data will be stored in an encrypted and password protected file, and will be retained indefinitely. Please note that if you fill out this survey on a public computer, you should remember to close your browser as soon as you have submitted your responses.

##### COMPENSATION

You will not be paid for taking part in this screening survey.

##### RIGHTS

Participation in research is completely voluntary. You are free to decline to take part in the project. You can decline to answer any questions and are free to stop taking part in the project at any time. Whether or not you partially or fully participate in the project, there will be no penalty to you or loss of benefits to which you are otherwise entitled.

##### QUESTIONS



If you have any questions or concerns about the research, you can contact me at [dianelam@berkeley.edu](mailto:dianelam@berkeley.edu).  
 If you have any questions or concerns about your rights and treatment as a research subject, you may contact the office of UC Berkeley's Committee for the Protection of Human Subjects, at 5106427461  
 Or [subjects@berkeley.edu](mailto:subjects@berkeley.edu).

#### CONSENT \*

If you agree to take part in the research, please print a copy of this page to keep for future reference, then click the checkbox below.

I have read the above information. By checking this box I am demonstrating my informed consent and agree to take part in this survey.

## Survey for Participation in Educational Research Study

1. How much do you agree with the following statements about your learning style? "I AM A \_\_\_\_\_ LEARNER"

Mark only one oval per row.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
VISUAL (prefer charts, graphs, symbols other than words)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
AURAL/AUDITORY (prefer spoken lessons and talking)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
READING/WRITING (prefer printed text)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
KINESTHETIC (prefer using my senses, direct concrete practice)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2. On average, how often do you use your science textbook for a course?

Mark only one oval per row.

Never	Less than once a week	Once a week	A few times a week	Everyday
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

3. Please briefly describe how you usually use your science textbooks for class (e.g. read carefully and take notes, skim the chapter, look at pictures).

4. How many of the following SCIENCE texts have you read in the past year?

Mark only one oval per row.

	None	Less than 5	Around 5	More than 10
Articles from primary literature (e.g. Nature, Cell)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Science-based magazines (e.g. Science, National Geographic)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Science Books (fiction or non-fiction)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Articles from credible online websites	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

5. Did you take and pass AP Biology in high school?

Mark only one oval.

- Yes  
 No

6. Have you completed a college-level biology course before this semester?

Mark only one oval.

- Yes  
 No

7 a. When was the last time you completed a biology course?

Please indicate the term and year:

Mark only one oval.

- Fall  
 Winter  
 Spring  
 Summer  
 2013  
 2012  
 2011  
 before 2011

b. What was the name of that biology course?

## Demographic Information

8. Are you a transfer student?

Mark only one oval.

- Yes  
 No

9. What is your current class standing?

Mark only one oval.

- Freshman  
 Sophomore  
 Junior  
 Senior

10. What is your major or intended major?

11. How old are you?

Mark only one oval.

- under 18 years old
- 18
- 19
- 20
- 21
- 22
- 23
- 24
- 25
- over 25 years old

12. What is your gender identity?

13. Please indicate your race or ethnicity.

Check all that apply.

- White
- Hispanic or Latino
- Black or African-American
- Native American, American Indian or Alaska Native
- Asian
- Native Hawaiian or Pacific Islander
- Other: \_\_\_\_\_

14. Please provide a first name and email address for us to contact you for scheduling.

## 10.2 Appendix B: Training PowerPoint Slides

### Training PowerPoint Slides

#### Self-Explanation Training

**Thank you for agreeing to take part in this experiment.**

Shortly, you will be presented with passages of text to study and asked to self-explain to aid your learning, but first you will receive training on how to make your self-explanations more effective for your learning.

#### Drawing Training

**Thank you for agreeing to take part in this experiment.**

Shortly, you will be presented with passages of text to study and asked to draw diagrams to aid your learning, but first you will receive training on how to make your drawings more effective for your learning.

#### What is Self-Explanation?

- Self-explanation is a study strategy that educational researchers have found enhances problem-solving and comprehension in learners.
- Note that a self-explanation is more than just a rewording or paraphrasing of what was stated in each sentence of text. It involves making connections between what you read and other sections of text or prior knowledge that you have related to the topic.

#### What is Drawing for Learning?

- Drawing is a study strategy that educational researchers have found enhances problem-solving and comprehension in learners.
- Note that drawing to learn is more than just illustrating exactly what was explicitly stated in each sentence of text. It involves drawing a cohesive picture of what was presented in the text as a whole, connecting all of the pieces to see how they fit together and fit with your prior knowledge.

#### Things to Consider When Self-Explaining

- Stop to self-explain whenever there is an opportunity to explain an idea to yourself.
  - These may be opportunities to learn something better.
- Explain to *yourself* and not to someone else in the room or someone you are imagining.
  - This will allow you to focus on your own understanding and learning, and not get distracted by how much another person may know.
- Don't worry about sounding eloquent or articulate, it is the self-explaining process rather than the outcome which will aid your learning.

#### Things to Consider When Drawing

- Stop to draw whenever there is an opportunity to illustrate an idea for yourself.
  - These may be opportunities to learn something better.
- Draw for *yourself* and not to someone else in the room or someone you are imagining.
  - This will allow you to focus on your own understanding and learning, and not get distracted by how much another person may know.
- Don't worry about being clear or neat: it is the drawing *process* rather than the outcome which will aid your learning.

### Things to Consider When Self-Explaining

A self-explanation	Not a self-explanation
• Elaborations on main ideas in the text	• A rewording or paraphrase
• Inferences about the content	• Defining words
• Connections between different sections of text.	• Monitoring statements (not focused on the content in the text)
• Connections to prior knowledge	
• Questions about the content	
• Predictions of what is to come next	

### Examples of Good and Poor Drawings

Good	Poor
• Summarizes and expands on key points	• No expansion on key points
• Functions as a cohesive model	• Disjointed and scattered
• Includes prior knowledge	• Vague descriptions
• Specific facts are included	• No specific details
• Important parts are labeled	• No labels

As a guide, you can ask yourself the following questions as you are self-explaining:

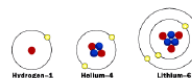
- 1. Main idea:** Do I understand the concept or main idea in the text? What kind of inferences can I make about these ideas?
- 2. Relate to other ideas:** How does this idea relate to other ideas in the text or to my prior knowledge?
- 3. My questions:** Does the self-explanation I have generated help to answer the questions that I am asking?

As a guide, you can ask yourself the following questions as you are drawing:

- 1. Coherent picture:** Is all of the information from the text present and connected together in my drawing?
- 2. Relate to other ideas:** Does my drawing include relevant references to other sections of text or to my prior knowledge?
- 3. Modelling:** Have I drawn ideas together in a way that forms a functional model of the system?

A model can be a diagram that abstracts and simplifies a system, by focusing on key features, to explain and predict scientific phenomena.

Ex. Bohr's model of atoms



On the following slides you will find examples of self-explanations generated by students while trying to understand a biology-related text.

We'll go through an example self-explanation and critique it together. Then, we'll end the training session by letting you practice generating your own self-explanation.

On the following slides you will find examples of diagrams drawn by students while trying to understand a biology-related text.

We'll go through an example diagram to explain and critique it together. Then, we'll end the training session by letting you practice generating your own diagram.

### Penguin Example

The internal temperature range of penguins is 37.8°C to 38.9°C (100°F-102°F). They also may shiver to generate additional heat. The dark plumage of a penguin's dorsal surface absorbs heat from the sun, which increases body temperature.

After reading this portion of text, one student generated this self-explanation:

"Okay, Penguins probably have a lot of tricks to get warm because they live in such a cold environment. I think 100-102 is a little warmer than humans who are at 98.6 degrees. Uh... And they shiver to get warm and their feathers are important for regulating their body temperature. Apparently the dorsal feathers help warm the penguin, which I guess would be on the penguin's head like the dorsal fin on a dolphin. I don't know. And the dark parts of a penguin's coat are important for keeping it warm, which would be similar to how a black car gets hotter than a white car on a sunny day. So, if their temperature drops below the normal range, they can try to bring it back up by shivering or by absorbing heat through their dark feathers."

The internal temperature range of penguins is 37.8°C to 38.9°C (100°F-102°F). They also may shiver to generate additional heat. The dark plumage of a penguin's dorsal surface absorbs heat from the sun, which increases body temperature.

- 1 "Okay, penguins probably have a lot of tricks to get warm because they live in such a cold environment.
- 2 I think 100-102 is a little warmer than humans who are at 98.6 degrees.
- 3 Uh... And they shiver to get warm and their feathers are important for regulating their body temperature.
- 4 Apparently the dorsal feathers help warm the penguin, which I guess would be on the penguin's head like the dorsal fin on a dolphin.
- 5 I don't know.
- 6 And the dark parts of a penguin's coat are important for keeping it warm, which would be similar to how a black car gets hotter than a white car on a sunny day.
- 7 So, if their temperature drops below the normal range, they can try to bring it back up by shivering or by absorbing heat through their dark feathers."

What parts of this example are good or poor self-explanations?

1. **Main idea:** Do I understand the concept or main idea in the text? What kind of inferences can I make about these ideas?
2. **Relate to other ideas:** How does this idea relate to other ideas in the text or to my prior knowledge?
3. **My questions:** Does my self-explanation I have generated help to answer the questions that I am asking?

The internal temperature range of penguins is 37.8°C to 38.9°C (100°F-102°F). They also may shiver to generate additional heat. The dark plumage of a penguin's dorsal surface absorbs heat from the sun, which increases body temperature.



What parts of this example make this a good or poor drawings?

1. **Coherent picture:** Is all of the information from the text present and connected together in my drawing?
2. **Relate to other ideas:** Does my drawing include relevant references to other sections of text or from my prior knowledge?
3. **Modelling:** Have I drawn ideas together in a way that forms a functional model of the system?

### Practice: Camel Example

A camel's eyes are protected by long, curly eyelashes that keep out sand. Thick eyebrows shield the eyes from the desert sun. The hump on a camel's back is not a storage for water, but for fat. It shrinks if the camel is starving.

1. **Main idea:** Do I understand the concept or main idea in the text? What kind of inferences can I make about these ideas?
2. **Relate to other ideas:** How does this idea relate to other ideas in the text or to my prior knowledge?
3. **My questions:** Does my self-explanation I have generated help to answer the questions that I am asking?

### Practice: Camel Example

A camel's eyes are protected by long, curly eyelashes that keep out sand. Thick eyebrows shield the eyes from the desert sun. The hump on a camel's back is not a storage for water, but for fat. It shrinks if the camel is starving.

Now please take a couple minutes to draw what you understand from the text, keeping in mind the techniques have discussed. You may create more than one drawing if you'd like.

1. **Coherent picture:** Is all of the information from the text present and connected together in my drawing?
2. **Relate to other ideas:** Does my drawing include relevant references to other sections of text or from my prior knowledge?
3. **Modelling:** Have I drawn ideas together in a way that forms a functional model of the system?

**Congratulations**

You have completed the self-explanation training session and you are now ready to continue with the experiment. During the experiment, you will be presented with the guiding questions we've been discussing to remind you and help you self-explain as you read the text.

Questions?

**Congratulations**

You have completed the drawing-to-learn training session and you are now ready to continue with the experiment. During the experiment, you will be presented with the guiding questions we've been discussing to remind you and help you to draw as you read the text.

Questions?

## 10.3 Appendix C: Cardiovascular System Text

---

### Cardiovascular System Text

#### 1. The Human Circulatory System

Human life depends on the distribution of oxygen, hormones, and nutrients to the cells in all parts of the body and on the removal of carbon dioxide and other wastes. These tasks are partially carried out by the circulatory system, which consists of the heart, and intricate network of blood vessels, and blood.

#### 2. The General Structure of the Heart

The septum divides the heart lengthwise into two sides. The right side pumps blood to the lungs, and the left side pumps blood to other part of the body. Each side of the heart is divided into an upper and a lower chamber. Each lower chamber is called a ventricle. Each upper chamber is called an atrium. In each side of the heart blood flows from the atrium to the ventricle.

#### 3. Blood Flow in the Heart

One-way valves separate these chambers and prevent blood from moving in the wrong direction. The atrioventricular valves (a-v) separate the atria from the ventricles. The a-v valve on the right side is the tricuspid valve, and the a-v valve on the left side is the bicuspid valve. Blood flows out of the ventricles. Two semilunar (s-l) valves separate the ventricles from the large vessels through which blood flows out of the heart. Each of the valves consists of flaps of tissue that open as blood is pumped out of the ventricles.

Blood returning to the heart, which has high concentration, or density, of carbon dioxide and a low concentration of oxygen, enters the right atrium. The atrium pumps it through the tricuspid valve into the right ventricle. Blood from the right ventricle flows through the semilunar valve into the pulmonary artery and then to the lungs. In the lungs, carbon dioxide leaves the circulating blood and oxygen enters it. The oxygenated blood returns to the left atrium of the heart. The oxygenated blood is then pumped through the bicuspid valve into the left ventricle. Blood from the left ventricle flows through the left semilunar valve, into the aorta, and then throughout the body.

#### 4. Phases of Muscle Contraction

The contractions which force blood through the semilunar valves and then to the lungs and throughout the body can be described as having two phases. In one phase, called systole, the ventricles contract and force blood into the arteries. In another phase, called diastole, the ventricles relax, and the blood flows in from the atria.

#### 5. Blood Vessels

The large, muscular vessels that carry blood away from the heart are called arteries. Blood travels through a network of smaller arteries, which in turn divide and form even smaller vessels called arterioles. The arterioles branch into a fan of tiny vessels called capillaries. De-oxygenated blood flows through capillaries that merge and form larger vessels called venules. Several venules in turn unite to form a vein, a large blood vessel that carries blood to the heart

#### 6. Structure of Blood Vessels



The thick walls of the arteries have three layers, an inner endothelium, a middle layer of smooth muscle, and an outer layer of connective tissue. Capillary walls are only one cell thick. The walls of veins, which are also composed of three layers, are thinner and less muscular than those of arteries.

#### 7. The Process of Diffusion

Diffusion is the process by which molecules spread from an area of greater concentration (or density) to an area of lesser concentration. When the concentration of the molecules of a substance is the same throughout a space, a state called equilibrium exists.

#### 8. The Process of Diffusion Through a Membrane

Molecules, including gases and nutrients, are in constant, random motion, traveling in straight lines until they hit something, then they rebound, traveling in a different direction. Once equilibrium is established, the random movement of molecules continues, and equilibrium is maintained. Some kinds of molecules can also pass through a membrane. They do this by moving between the molecules that make up the membrane. If a substance can pass through a membrane, the membrane is permeable to it.

#### 9. Function of Valves

Valves prevent the blood from moving backward or downward. These valves allow blood to flow in only one direction through the veins.

#### 10. The Two Primary Subsystems of Circulation

The two primary subsystems of circulation are pulmonary circulation, in which the blood travels between the heart and lungs, and systemic circulation, in which the blood travels between the heart and all other body tissues.

#### 11. Details of the Primary Subsystems of Circulation

Pulmonary Circulation is the movement of blood from the heart to the lungs and back to the heart. The pulmonary artery is the only artery that carries deoxygenated blood. At the lungs the pulmonary artery divides into two smaller arteries one leading to each lung. These arteries branch into arterioles and then into capillaries. Oxygenated blood then flows into venules, which merge into the pulmonary veins that lead to the left atrium of the heart. The pulmonary veins are the only veins that carry oxygenated blood. Systemic Circulation is the movement of blood to all parts of the body, except the lungs, and then back to the heart. Blood from the left ventricle enters systemic circulation through the aorta.

## 10.4 Appendix D: Free-Response Questions

### Free-Response Questions on Formal Assessments

† Indicates questions that were not used in the analyses because of unreliability

#### Question Set A

S	P	F	
			<b>Definition Questions:</b>
1		1	Artery
1		1	Capillary
2		1	Pulmonary circulation
1		1	Septum
2		1	Ventricle
			<b>Explicit Questions:</b>
2			Which blood vessels carry oxygenated blood? Where is blood oxygenated?( <i>Absorbed by Pulmonary Circulation Definition</i> )
1			Where does the right ventricle pump blood to?
1			Where does blood go to after it's oxygenated?
2			Why is blood vital to human life? That is, what is its primary purpose?
			<b>Implicit Questions:</b>
			Do veins carry blood in any particular direction? If so, which direction?( <i>Absorbed by Vein Definition</i> )
1			Why are capillary walls only one cell thick? †Does the blood change in any way as it passes through the heart? If so, how?
2			What keeps the blood flowing in the given direction when it leaves the heart?
			<b>Mental Model Questions:</b>
1			What is the consequence at the cellular level of having a hole in the septum?
1			Why is your right ventricle less muscular than your left ventricle?
1			Why do we sometimes refer to the heart as a “doublepump”? †The artery that carries blood from the right side of the heart to the lungs (the pulmonary artery) carries about the same amount of blood as the artery that carries blood from the left side of the heart to the rest of the body (aorta). Why do they carry the same amount of blood?
7	4	13	

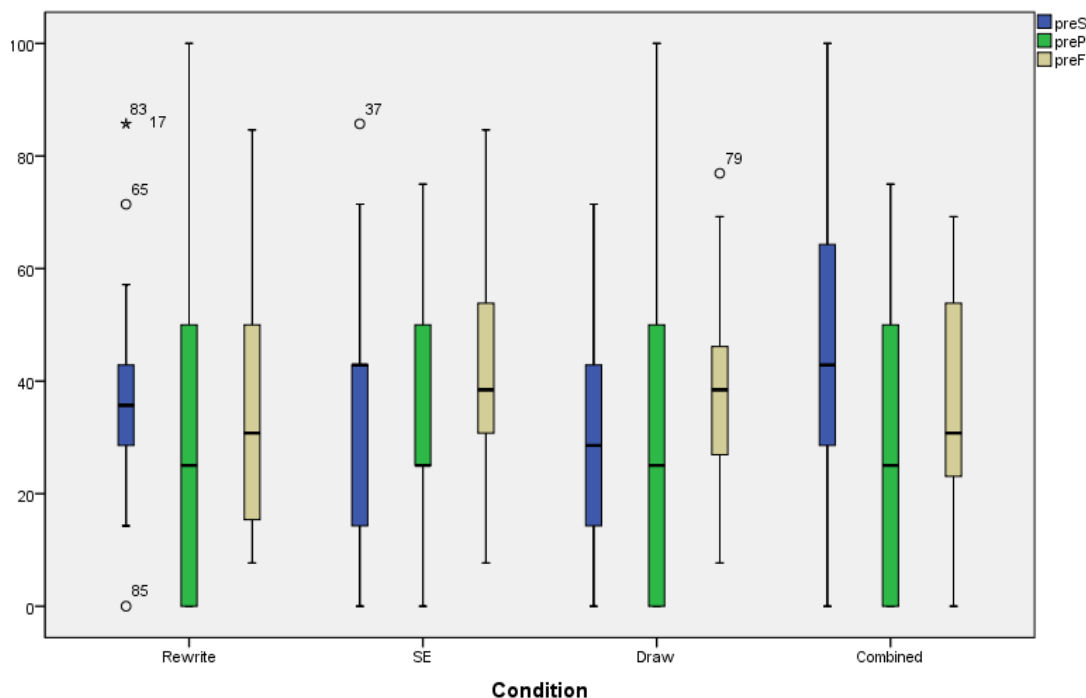
## Question Set B

S	P	F	
			<b>Definition Questions:</b>
2	1		Atrium †Heart
2	1		Systemic circulation
1	1		Valve
1	1		Vein
			<b>Explicit Questions:</b>
2			Which blood vessels carry deoxygenated blood?
1			Where does the left ventricle pump blood to?
1			Where does blood go after it delivers oxygen to body cells? †What does the heart do? <i>(Absorbed by Heart definition)</i> Why are there valves in the heart? <i>(Absorbed by Valve Definition)</i>
			<b>Implicit Questions:</b>
1	2		In which kind of blood vessels is the blood pressure the lowest? Why?
1	1		Is there an artery that carries deoxygenated blood? Please explain.
	1		Why do vessels get increasingly smaller as they get close to the body cells, and get increasingly larger as they get nearer to the heart?
			<b>Mental Model Questions:</b>
	1		Why does the circulatory system need an atrium? Why can't blood flow straight through each ventricle in a pipe-like fashion?
	1		Why can't the blood go straight to the lungs after it comes back from the body? In other words, why does the blood need to go to the heart first before it goes to the lungs?
	1		What would happen if the valves between the atria and the ventricles are stuck open and wouldn't close?
8	4	11	

## 10.5 Appendix E: Pre-test MANOVA Assumptions

Seven outliers were identified via boxplots (Figure ), but were not removed from the dataset because we believed this data to be representative of the groups and would not materially affect the results. Not all data were normally distributed for each group, but we proceeded with the MANOVA since the analysis is robust to deviations from normality. There was no multicollinearity ( $VIFs < 2.0$ ). A linear relationship between pretest structure, pathway, and function scores for each group, as assessed by scatterplots. There were no multivariate outliers in the data, as assessed by Mahalanobis distance ( $p < 0.001$ ).

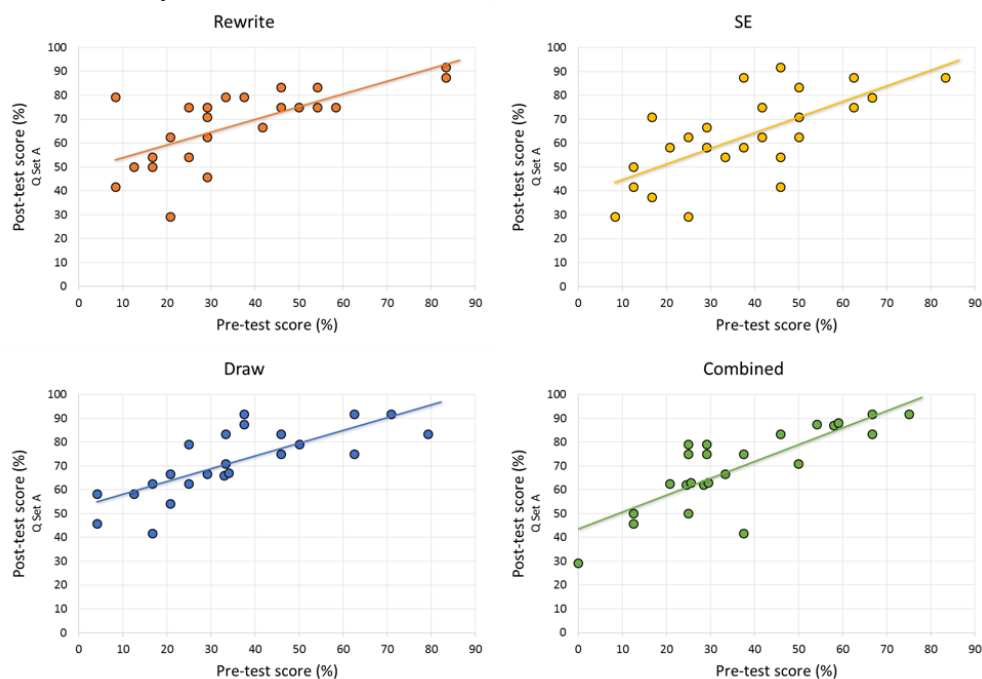
There was not a homogeneity of variance-covariances matrices, but since the groups have similar n sizes, we proceeded with the MANOVA. There was homogeneity of variances, as assessed by Levene's Test of Homogeneity of Variances ( $p > 0.05$ )



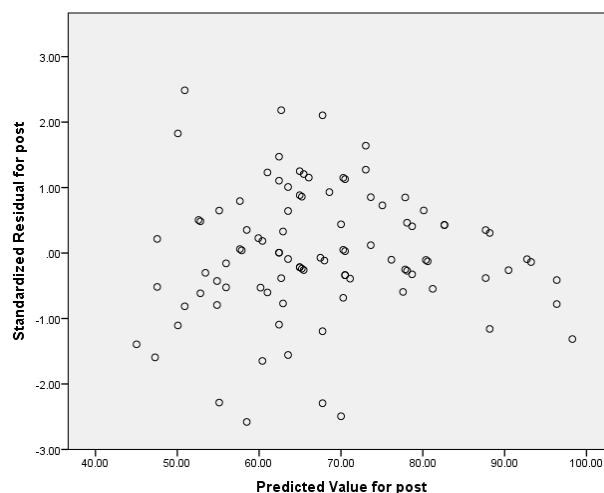
**Figure E1:** Boxplot of pre-test structure, pathway, and function scores.

## 10.6 Appendix F: Post-test QSA ANCOVA Assumptions

Regression diagnostics determined that an ANCOVA was appropriate for these data. There was a linear relationship between pre- and post-test QSA scores in each condition, as assessed by visual inspection of scatterplots and trend lines (Figure F1). The interaction term between pre- and post-test scores was not statistically significant,  $F(3,88) = 0.512$ ,  $p = 0.675$ , indicating homogeneity of regression slopes. Furthermore, standardized residuals for the groups and the overall model were normally distributed, as assessed by Shapiro-Wilk's test ( $p > 0.05$ ). A visual inspection of a scatterplot of predicted Post-test scores and standardized residuals indicated homoscedasticity of variances (Figure F2), and Levene's test indicated homogeneity of variance ( $p=0.427$ ). No standardized residuals were greater than three standard deviations, thus we did not consider any cases as outliers.



**Figure F1:** Scatterplots and trend lines for pre-test vs. post-test QSA scores, by condition, show a linear relationship between the variables.



**Figure F2:** Scatter plot of predicted post-test QSA scores and standardized residuals shows homoscedasticity of variances.

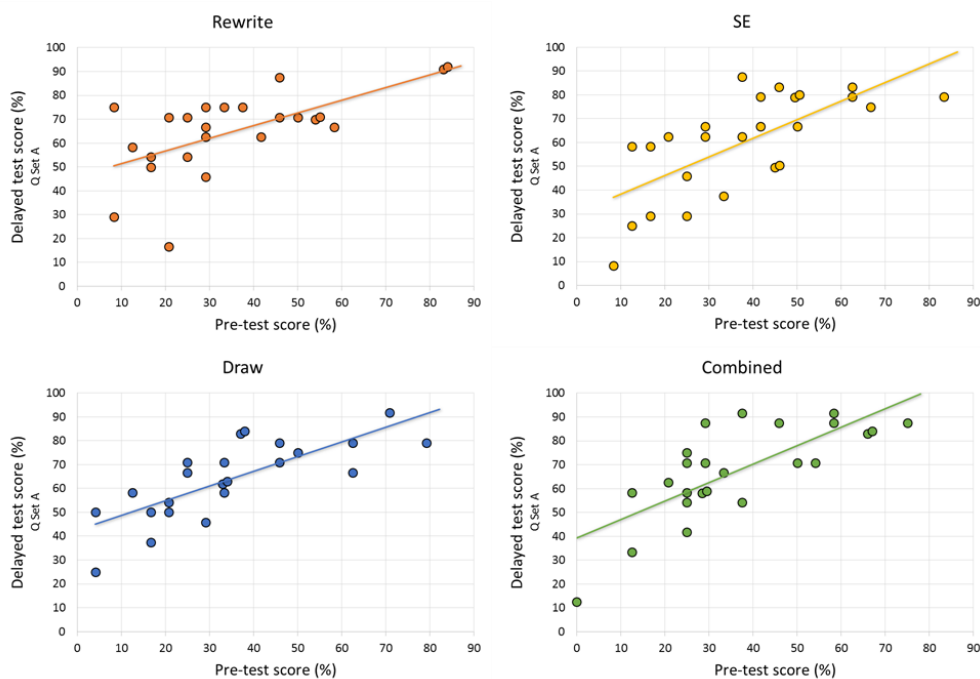
Adjusted mean scores for post-test QSA scores are presented in Table F1 below. These scores were adjusted using overall pre-test scores or SPF sub-scores as covariates in the ANCOVA.

**Table F1:** Adjusted means (and standard errors) of post-test QSA scores, using pre-test scores as covariates.

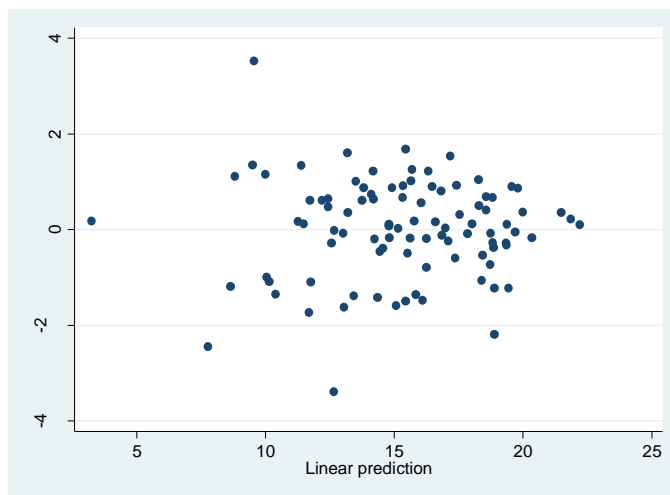
Condition	N	Adjusted Post-test QSA Scores (%)			
		Overall	Structures	Pathways	Functions
Rewrite	24	67.82 (2.32)	88.77 (3.23)	60.00 (4.45)	59.27 (2.68)
SE	25	61.93 (2.28)	71.77 (3.17)	57.21 (4.36)	57.65 (2.63)
Draw	23	72.26 (2.37)	90.18 (3.32)	72.67 (4.54)	61.46 (2.73)
Combined	24	69.25 (2.32)	82.19 (3.26)	69.10 (4.46)	63.44 (2.68)

## 10.7 Appendix G: Delayed-test QSA ANCOVA Assumptions

There was a linear relationship between pre- and delayed-test QSA scores in each condition, as assessed by visual inspection of scatterplots and trend lines (Figure G1). The interaction term between pre- and delayed-test scores was not statistically significant,  $F(3,88) = 0.709$ ,  $p = 0.549$ , indicating homogeneity of regression slopes. Standardized residuals for the groups and the overall model were normally distributed, as assessed by Shapiro-Wilk's tests ( $p > 0.05$ ). A visual inspection of a scatterplot of predicted delayed-test scores and standardized residuals indicated homoscedasticity of variances (Figure G2), and Levene's test indicated homogeneity of variance ( $p=0.203$ ). No standardized residuals were greater than three standard deviations, thus we did not consider any cases as outliers.



**Figure G1:** Scatterplots and trend lines for pre-test vs. delayed-test QSA scores, by condition, show a linear relationship between the variables.



**Figure G2:** Scatter plot of predicted delayed-test QSA scores and standardized residuals shows homoscedasticity of variances

Adjusted mean scores for delayed-test QSA scores are presented in Table G1 below. These scores were adjusted using overall pre-test scores or SPF sub-scores as covariates in the ANCOVA.

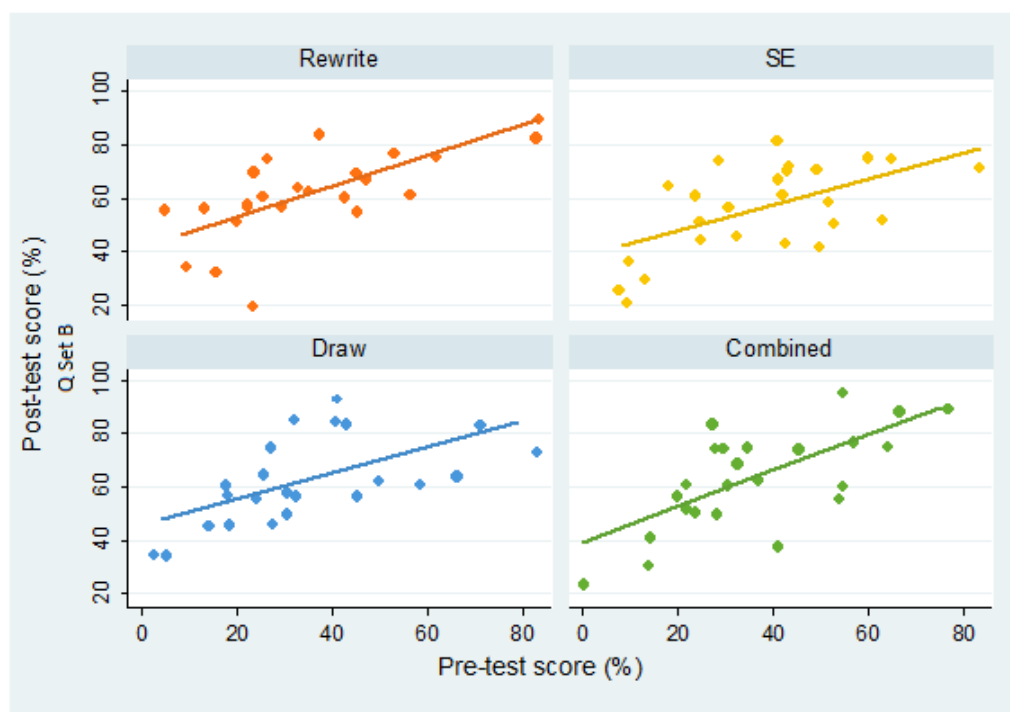
**Table G1:** Adjusted means (and standard errors) of delayed-test QSA scores, using pre-test scores as covariates.

Condition	N	Adjusted Delayed-test QSA Scores (%)			
		Overall	Structures	Pathways	Functions
Rewrite	24	65.42 (2.75)	83.44 (4.47)	61.04 (4.93)	57.44 (2.61)
SE	25	58.99 (2.69)	69.01 (4.38)	48.21 (4.84)	56.58 (2.56)
Draw	23	64.93 (2.81)	78.12 (4.59)	62.89 (5.04)	57.42 (2.67)
Combined	24	67.33 (2.75)	76.96 (4.52)	62.85 (4.95)	64.49 (2.61)

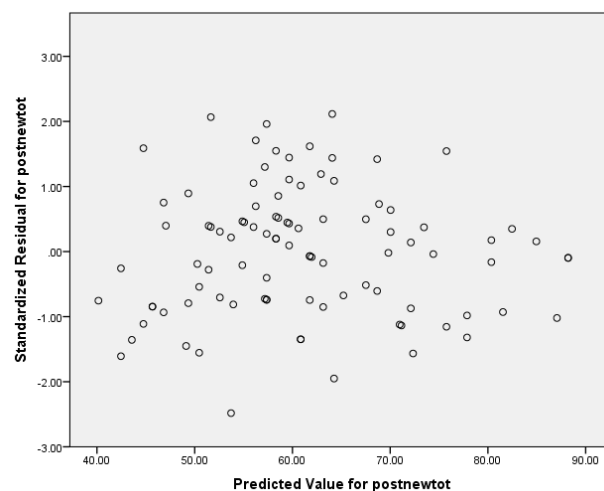


## 10.8 Appendix H: Post-test QSB ANCOVA Assumptions

There was a linear relationship between pre and post-test scores for Question set B in each condition (Figure H1). The interaction term between pre- and post-test scores was not statistically significant,  $F(3,88) = 0.395$ ,  $p = 0.757$ , indicating homogeneity of regression slopes. Furthermore, standardized residuals for the groups and the overall model were normally distributed, as assessed by Shapiro-Wilk's test ( $p \leq 0.05$ ). A visual inspection of a scatterplot of predicted Post-test scores and standardized residuals indicated homoscedasticity of variances (Figure H2). Levene's test indicated homogeneity of variance ( $p = 0.297$ ). No standardized residuals were greater than three standard deviations, thus we did not consider any cases as outliers.



**Figure H1:** Scatterplots and trend lines for pre-test vs. post-test QSB scores, by condition, show a linear relationship between the variables.



**Figure H2:** Scatter plot of predicted post-test QSB scores and standardized residuals shows homoscedasticity of variances

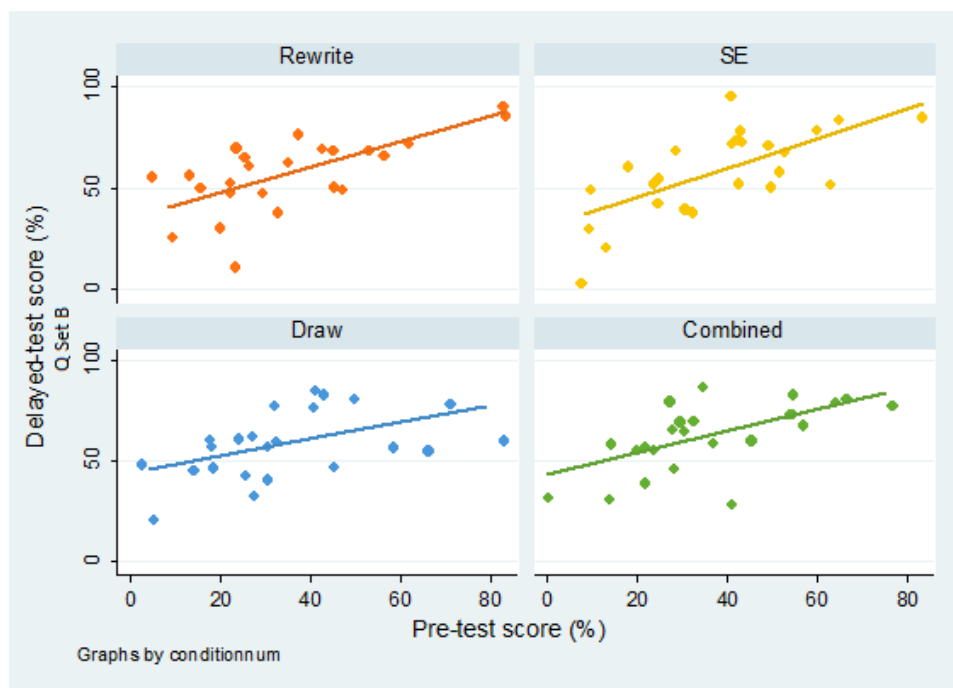
Adjusted mean scores for post-test QSB scores are presented in Table H1 below. These scores were adjusted using overall pre-test scores or SPF sub-scores as covariates in the ANCOVA.

**Table H1:** Adjusted means (and standard errors) of post-test QSB scores, using pre-test scores as covariates.

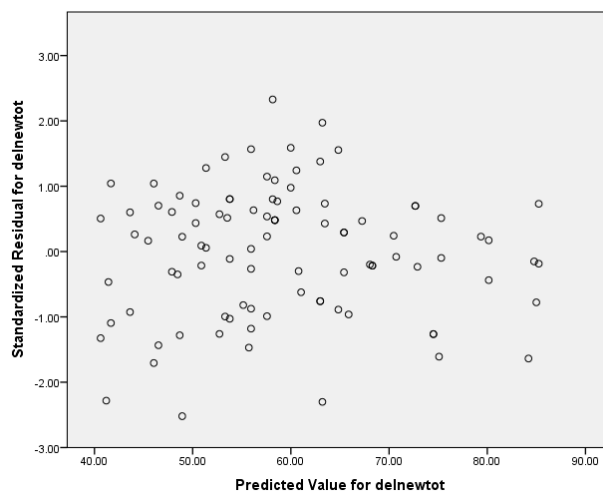
Condition	N	Adjusted Post-test QSB Scores (%)			
		Overall	Structures	Pathways	Functions
Rewrite	24	62.22 (2.63)	74.59 (3.36)	67.60 (3.91)	51.33 (3.42)
SE	25	55.55 (2.58)	64.95 (3.29)	64.58 (3.84)	45.73 (3.36)
Draw	23	63.38 (2.69)	76.23 (3.45)	72.48 (3.99)	50.28 (3.49)
Combined	24	63.56 (2.63)	71.16 (3.39)	74.84 (3.92)	53.98 (3.42)

## 10.9 Appendix I: Delayed-test QSB ANCOVA Assumptions

There was a linear relationship between pre and delayed-test scores for Question set B in each condition (Figure I1). The interaction term between pre- and delayed-test scores was not statistically significant,  $F(3,88) = 0.765$ ,  $p = 0.517$ , indicating homogeneity of regression slopes. Furthermore, standardized residuals for the groups and the overall model were normally distributed, as assessed by Shapiro-Wilk's test ( $p \leq 0.05$ ). A visual inspection of a scatterplot of predicted Delayed-test scores and standardized residuals indicated homoscedasticity of variances (Figure I2). Levene's test indicated homogeneity of variance ( $p = 0.436$ ). No standardized residuals were greater than three standard deviations, thus we did not consider any cases as outliers.



**Figure I1:** Scatterplots and trend lines for pre-test vs. delayed-test QSB scores, by condition, show a linear relationship between the variables.



**Figure I2:** Scatter plot of predicted delayed-test QSB scores and standardized residuals shows homoscedasticity of variances

Adjusted mean scores for delayed-test QSB scores are presented in Table II below. These scores were adjusted using overall pre-test scores or SPF sub-scores as covariates in the ANCOVA.

**Table II:** Adjusted means (and standard errors) of delayed-test QSB scores, using pre-test scores as covariates.

Condition	N	Adjusted Delayed-test QSB Scores (%)			
		Overall	Structures	Pathways	Functions
Rewrite	24	57.89 (2.91)	67.30 (4.04)	62.39 (4.02)	49.58 (3.43)
SE	25	57.41 (2.85)	61.46 (3.96)	57.58 (3.94)	54.59 (3.37)
Draw	23	59.26 (2.97)	65.40 (4.14)	65.95 (4.10)	51.82 (3.50)
Combined	24	62.48 (2.91)	68.51 (4.07)	70.68 (4.03)	55.26 (3.43)

### 10.10 Appendix J: Correlation Matrix of Time on Task and Sub-scores

		Rewrite Group		SE group		Draw group		Combined group			Totals across groups			
		Rewrite time	Total time	SE time	Total time	Draw time	Total time	SE time	Draw time	Total time	SE time	Draw time	Strategy time	
	Total S	-0.132	0.025	0.170	0.127	0.394	<b>*0.442</b>	0.138	-0.307	-0.289	0.116	0.034	<b>*0.222</b>	
	Total P	-0.066	-0.019	0.235	0.276	-0.218	-0.155	0.315	-0.066	-0.084	0.238	-0.109	0.139	
	Total F	-0.155	-0.084	0.197	0.160	0.231	0.288	0.275	-0.213	-0.214	0.223	0.005	0.109	
Post-Test	QSA	Overall	-0.066	-0.037	0.346	0.269	0.142	0.215	0.322	-0.368	-0.296	<b>*0.307</b>	-0.031	0.185
		S	-0.070	0.000	0.179	0.041	0.112	0.148	0.220	<b>*-0.455</b>	-0.231	0.151	-0.056	0.206
		P	-0.027	0.021	0.327	0.383	-0.058	-0.013	0.127	-0.264	-0.301	0.217	-0.059	0.145
		F	-0.063	-0.062	0.335	0.265	0.180	0.257	0.345	-0.246	-0.230	<b>*0.329</b>	0.000	0.127
	QSB	Overall	-0.205	-0.064	0.185	0.188	0.186	0.246	0.285	-0.189	-0.162	0.200	0.004	0.179
		S	-0.093	0.058	0.192	0.155	0.050	0.127	0.133	-0.267	-0.258	0.130	-0.061	0.156
		P	-0.205	-0.152	0.267	0.351	-0.271	-0.231	0.358	0.060	0.140	0.282	-0.113	0.114
		F	-0.248	-0.115	0.088	0.087	0.360	0.391	0.284	-0.170	-0.157	0.165	0.085	0.163
	QSA&B	Overall	-0.141	-0.053	0.288	0.245	0.169	0.238	0.310	-0.284	-0.233	0.266	-0.014	0.190
		S	-0.099	0.043	0.209	0.110	0.094	0.167	0.186	-0.380	-0.266	0.154	-0.066	0.202
		P	-0.104	-0.052	0.318	0.390	-0.181	-0.137	0.249	-0.132	-0.116	0.262	-0.091	0.138
		F	-0.152	-0.091	0.250	0.206	0.289	0.347	0.329	-0.218	-0.203	0.270	0.044	0.154
Delayed-	QS	Overall	-0.040	0.050	0.118	0.081	0.045	0.125	0.299	-0.280	-0.254	0.172	-0.118	0.121
		S	0.046	0.153	0.048	0.025	0.393	0.427	0.063	-0.210	-0.262	0.024	0.085	0.185

	P	0.030	0.077	0.167	0.161	-0.214	-0.173	0.369	0.007	-0.018	0.224	-0.075	0.136
	F	-0.106	-0.038	0.119	0.065	-0.083	-0.005	0.331	-0.347	-0.263	0.207	-0.209	0.030
QSB	Overall	-0.232	-0.125	0.173	0.189	<b>*0.467</b>	<b>*0.498</b>	0.075	-0.043	-0.176	0.118	0.134	0.143
	S	-0.287	-0.134	0.187	0.209	<b>**0.564</b>	<b>**0.569</b>	0.084	-0.166	-0.240	0.114	0.117	0.205
	P	-0.097	-0.070	0.080	0.106	-0.203	-0.102	0.248	0.019	-0.064	0.113	-0.143	0.091
	F	-0.190	-0.108	0.149	0.151	0.385	0.387	-0.013	0.049	-0.115	0.089	0.195	0.065
QS A&B	Overall	-0.138	-0.037	0.150	0.138	0.273	0.337	0.208	-0.183	-0.229	0.153	-0.004	0.137
	S	-0.137	0.007	0.127	0.127	<b>*0.521</b>	<b>*0.542</b>	0.080	-0.205	-0.274	0.075	0.110	0.209
	P	-0.029	0.009	0.134	0.142	-0.230	-0.154	0.333	0.012	-0.038	0.183	-0.112	0.121
	F	-0.152	-0.074	0.143	0.114	0.151	0.201	0.202	-0.196	-0.216	0.165	-0.038	0.050
Short-term MM gain		-0.065	0.180	0.198	0.387	0.296	0.274	-0.150	0.218	0.130	0.019	0.300	<b>***0.344</b>
Long-term MM gain		-0.217	-0.003	0.146	0.207	0.418	0.370	-0.125	0.189	0.145	-0.027	<b>*0.333</b>	<b>**0.315</b>

\*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$

### 10.11 Appendix K: Correlation Analyses of Full Self-explanation Data

Presented in Table K1 and Table K2 are correlation matrices for codes for self-explanation and formal assessment scores, including the data from outliers that were dropped for correlation results presented in Table 14 and Table 15. Stronger correlations and more correlation values that are statistically significant were found when the outliers were included in the dataset. The outliers that were dropped were participants who generated a substantially greater number of self-explanations, and also performed relatively high on the post and delayed tests. All patterns in these correlations are the same as with the dataset without outliers, but all correlation values are stronger.

**Table K1:** Correlation matrix of formal assessment sub-scores and DMM scores versus total self-explanation and self-explanation formats for full self-explanation dataset (outliers kept), pooled from the SE and Combined groups.

Scores	Total No. of SEs	Self-explanation Formats		
		Goal-Driven	Elaborative	Noticing Coherence
<b>Total SPF</b>	<b>0.35*</b>	<b>0.48***</b>	0.26	<b>0.29*</b>
<b>Total S</b>	0.26	<b>0.40**</b>	0.20	0.21
<b>Total P</b>	<b>0.31*</b>	<b>0.34*</b>	0.22	0.25
<b>Total F</b>	<b>0.37**</b>	<b>0.51***</b>	0.26	<b>0.30*</b>
<b>DMM</b>	0.08	0.07	0.06	0.08

\* $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\* $p < 0.001$

**Table K2:** Correlation matrix of formal assessment sub-scores and DMM scores versus mentions of SPF categories in self-explanations for full self-explanation dataset from the SE and Combined groups.

Scores	Mention of SPF categories in self-explanations		
	S	P	F
<b>Total SPF</b>	<b>0.34*</b>	<b>0.36**</b>	<b>0.42**</b>
<b>Total S</b>	<b>0.30*</b>	0.27	<b>0.29*</b>
<b>Total P</b>	<b>0.35*</b>	<b>0.35*</b>	<b>0.35*</b>
<b>Total F</b>	<b>0.29*</b>	<b>0.37*</b>	<b>0.46*</b>
<b>DMM</b>	0.09	0.09	0.05