

The University of Maine

DigitalCommons@UMaine

Electronic Theses and Dissertations

Fogler Library

Summer 8-21-2020

How Students Communicate Knowledge: Written Versus Drawn Responses to Formative Assessment Questions in an Introductory Undergraduate Marine Science Course

Christina Siddons

University of Maine, christina.siddons@maine.edu

Follow this and additional works at: <https://digitalcommons.library.umaine.edu/etd>



Part of the [Marine Biology Commons](#), and the [Science and Mathematics Education Commons](#)

Recommended Citation

Siddons, Christina, "How Students Communicate Knowledge: Written Versus Drawn Responses to Formative Assessment Questions in an Introductory Undergraduate Marine Science Course" (2020). *Electronic Theses and Dissertations*. 3227.

<https://digitalcommons.library.umaine.edu/etd/3227>

This Open-Access Thesis is brought to you for free and open access by DigitalCommons@UMaine. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of DigitalCommons@UMaine. For more information, please contact um.library.technical.services@maine.edu.

**HOW STUDENTS COMMUNICATE KNOWLEDGE:
WRITTEN VERSUS DRAWN RESPONSES TO
FORMATIVE ASSESSMENT QUESTIONS IN
AN INTRODUCTORY UNDERGRADUATE
MARINE SCIENCE COURSE**

By

Christina A. Siddons

B.S. Biology, Cornell University, 2010

A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

(in Teaching)

The Graduate School

University of Maine

August 2020

Advisory Committee:

Sara Lindsay, Associate Professor of Marine Sciences, Advisor

Asli Sezen-Barrie, Associate Professor of Curriculum, Assessment, and Instruction

Natasha M. Speer, Associate Professor of Mathematics Education

© 2020 Christina A. Siddons

All Rights Reserved

**HOW STUDENTS COMMUNICATE KNOWLEDGE:
WRITTEN VERSUS DRAWN RESPONSES TO
FORMATIVE ASSESSMENT QUESTIONS IN
AN INTRODUCTORY UNDERGRADUATE
MARINE SCIENCE COURSE**

By Christina A. Siddons

Thesis advisor: Dr. Sara Lindsay

An Abstract of the Thesis Presented
in Partial Fulfillment of the Requirements for the
Degree of Master of Science
(in Teaching)
August 2020

Undergraduate science education suffers from a lack of concrete instructional strategies that address real-world postgraduate skills such as visual literacy and science communication. Research within marine science education especially lags behind other, more well-researched fields such as physics or mathematics education, both of which have extensive literature addressing specific instructional strategies that instructors can implement in the classroom. Undergraduate marine science programs overlap with content areas from chemistry, physics, and biology, and provide a rich opportunity for examining how to include more authentic educational experiences in an undergraduate classroom. However, the types of assessments that are typically employed tend to encourage practices such as rote memorization and fact-recall, as assessed by lengthy multiple-choice quizzes and exams. Such assessments have come under scrutiny as professionals and educators alike call for undergraduate instruction to more closely align with

actual scientific practice. This study assessed a drawing-to-learn strategy in a marine science classroom to determine if opportunities for students to utilize diagramming and drawing during formative assessments translated into greater depth of information and understanding obtained from their responses.

Three different years of student cohorts enrolled in an introductory marine science course at a public university in the Northeastern United States that focused on comparative anatomy and evolution of marine phyla were given formative assessment “notecard questions” throughout the semester-long course from 2017 – 2019. A prompt regarding the close linkages between circulatory and respiratory systems – which exemplified core concepts from guiding instructional documents, as well as addressed specific course goals – was examined in detail, with responses from 2017 and 2018 comprising of traditional written answers, whereas 2019 responses were drawn. Notecards were coded for a variety of holistic and specific parameters to determine the detail of response, whether alternative conceptions were present, and expertise of response, comparing written responses to drawn.

Results indicated that drawn responses tended to capture more core ideas (“Key Concepts”) out of three identified and greater depth of detail than written alone. In particular, drawn responses captured specific structures such as the heart (58.2% of responses) and lungs/gills (84.8% of responses) as compared to only 7.3% ($\chi^2 = 73.08$, $df = 1$, $p < 0.001$) and 43.8% ($\chi^2 = 38.26$, $df = 1$, $p < 0.001$) of written responses, respectively. Certain Key Concepts also seemed to be more easily depicted in drawn form than written, such as the idea of circulatory – respiratory integration. Interestingly, although both response categories had alternative conceptions present, certain alternative conception codes that were more frequent in

drawn responses required a higher threshold of knowledge for students to demonstrate before such a code could be invoked.

Taken together, the results from this study reveal that strategically incorporating drawing-to-learn opportunities in the undergraduate marine science classroom can provide instructors with more insight into student knowledge than writing alone. Future research can build upon the approaches taken in this study to implement more scaffolded approaches to drawing and diagramming in order to meet the challenges of providing authentic scientific learning opportunities.

ACKNOWLEDGMENTS

I have been so lucky to have three incredible role models to serve on my committee for this thesis project, and am deeply indebted to them for the help and support they have given me throughout the past two years. I would like to first extend my deepest gratitude to my advisor, Dr. Sara Lindsay, for her endless support and guidance during my time in the MST program and throughout this research project. I so appreciate the academic and personal support and caring you have shown me, and am grateful I had the opportunity to work with you. Thank you also to Drs. Natasha Speer and Asli Sezen-Barrie, both of whom have provided much-valued insights, comments, and support throughout my thesis research and writing; I am grateful for both of your expertise. Thank you also to Michael Wittman, who has provided thoughtful insights and guidance for nitty-gritty details and important big-picture questions, both through formal class work and informal discussions. A big thanks to Susan McKay for your support and encouragement while I was considering and applying to the MST program, and for recognizing my potential as a future teacher when I wasn't sure about it myself. And a million thanks to my MST/NSF cohorts: you are all incredibly smart, funny, and generous, and I learned so much from all of you. Thanks for keeping me sane and teaching me physics- I can't wait to see what our futures hold!

Finally, I owe an enormous thanks to my husband, Stu Siddons, whose constant patience and encouragement throughout the long hours, frustrations, victories, and failures has kept my head on straight throughout this entire process. I could not have done this program without you by my side, and for that I am so full of gratitude and love.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	III
LIST OF TABLES	VIII
LIST OF FIGURES	IX
LIST OF ABBREVIATIONS	XI
CHAPTER 1: INTRODUCTION	1
1.1 The need for marine science in science curricula	4
1.2 Marine science, meaningful science learning, and instruction	5
1.3 The research deficit regarding marine science education	7
1.4 Drawing-to-learn in the marine science classroom	8
1.5 Visual literacy and STEM education.....	11
1.5.1 Visual literacy	11
1.5.2 Discursive fluency and addressing visual communication skills	12
1.5.3 Visual literacy in science	13
1.6 Drawing-to-learn: strategies and instructional goals	15
1.6.1 Defining ‘drawing’ and ‘drawing-to-learn’	15
1.6.2 Drawing-to-learn framework in STEM.....	16
1.7 Information Processing Theory	19
1.7.1 Dual-coding and generative learning theories	20
1.7.2 Mayer’s cognitive theory of multimedia learning	21
1.8 Cognitive processes activated by imaged-based learning	22
1.8.1 Selection – Organization – Integration	22

1.8.2 Metacognitive skills	25
1.8.3 Visual literacy and cognitive theory	26
1.9 Implementing meaningful learning strategies: visual literacy and drawing-to-learn	27
1.9.1 Diversifying assessment: the value of question type	29
CHAPTER 2: METHODS	32
2.1 Data collection	32
2.2 Grounded theory and multimedia discourse analysis	33
2.3 Determining expert versus novice responses	34
2.4 Codebook construction	37
2.4.1 Key Concept codes	39
2.4.2 Degree of Visualization and Depth of Drawing	41
2.4.3 Representations	42
2.4.4 Coherence	44
2.4.5 Expert Score	44
2.4.6 Alternative Conceptions	45
2.5 Data Analysis	49
2.5.1 Sample pooling	49
2.5.2 Frequency calculations	49
2.5.3 Chi-square analysis	50
CHAPTER 3: RESULTS	51
3.1 Student communication of knowledge	51
3.1.1 Key Concept codes	52
3.1.2 Representation codes	58

3.2 Assessing novice versus expert responses	59
3.2.1 Expert Score breakdown by response type	60
3.2.2 Key Concept co-occurrence with Expert Score.....	62
3.2.3 Expert Score co-occurrence with drawing-specific codes	65
3.3 Alternative conceptions in novice responses	67
CHAPTER 4: DISCUSSION	71
4.1 Student knowledge: drawing may facilitate communication of information.....	72
4.1.1 Visualization appears to facilitate cognitive coding and expression of certain Key Concept information	73
4.1.2 Depicting countercurrent exchange accurately is more difficult than other Key Concepts.....	77
4.1.3 Finer-grained resolution of structures are more easily represented visually than verbally	81
4.2 Demonstration of expert thinking: drawn versus written.....	83
4.2.1 What differentiates expert thinking?.....	83
4.2.2 Drawn responses.....	84
4.3 Alternative Conceptions and their impact on determining expertise	86
4.3.1 Certain alternative conceptions lend themselves to appearing in written form.....	87
4.3.2 Earning the “Capillaries Absent” Alternative Conception code itself requires a level of expertise	88
4.4 Limitations and contributions of current study.....	90
4.4.1 Scaffolding drawing opportunities during instruction may remove the barrier to entry	90

4.4.2 Assessment type may affect student effort in response	93
4.4.3 Notecards alone cannot capture the total picture of student understanding	94
4.4.4 Contributions to marine science education	95
CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS	98
REFERENCES.....	102
APPENDIX A	109
BIOGRAPHY OF THE AUTHOR.....	121

LIST OF TABLES

Table 2.3.1. Possible expert written responses.	35
Table 2.4.1 Description of Expert Score (ES) categories.....	45
Table 3.3.1. Raw count data for Alternative Conception (AC) codes by Expert Scores (ES) per response category.....	69
Table A.1. Codes developed and their descriptions.....	109

LIST OF FIGURES

Figure 1.4.1 Examples of common diagrams found in marine science instruction and publications.	9
Figure 1.6.1 Examples of diagrams depicting the linkage between circulatory and respiratory systems.....	18
Figure 2.3.1. Possible expert responses for 2019 drawn prompt.	36
Figure 2.4.1 Key Concept student examples.....	41
Figure 2.4.2 Student examples of various Representations codes.....	43
Figure 2.4.3 Student examples of the “Circulatory Inaccuracy” Alternative Conception (AC) code.....	46
Figure 2.4.4. Examples of coded 2017 student written responses.....	47
Figure 2.4.5. Examples of coded 2019 student drawn responses.	48
Figure 3.1.1. Number of Key Concepts coded by response type.....	53
Figure 3.1.2. Type of Key Concept code by response type.	54
Figure 3.1.3. Use of Evidence codes per Key Concept for drawn responses.	56
Figure 3.1.4. Card count data for co-occurring Key Concepts.	56
Figure 3.1.5. Representation codes frequencies by response type.....	59
Figure 3.2.1. Frequency of Expert Scores (ES) between the two response categories.	61
Figure 3.2.2. Sample of Expert Score (ES) of 0 in a drawn response.	62
Figure 3.2.3. Frequencies of number of Key Concepts (KC) across Expert Scores (ES).....	63
Figure 3.2.4. Frequency of each Key Concept (KC) code type by response category and Expert Score (ES).....	64
Figure 3.2.5. Depth of Drawing (DD) scores per Expert Score category.	66

Figure 3.3.1. Frequency of Alternative Conception (AC) codes.	68
Figure 3.3.2. Alternative Conception code frequency by response category amongst novice responses.....	70
Figure 4.1.1 Circulatory – respiratory integration (CRI) Key Concept communicated in drawn form.....	74
Figure 4.1.2 Visual examples of Key Concepts from lecture slides.....	76
Figure 4.1.3. Examples of mass transport diagrams from lecture slides.....	79
Figure 4.1.4. Examples of countercurrent exchange diagrams from lecture slides.....	80
Figure 4.1.5. Illustrative student examples of “System-Level” Representation code versus finer-grained structural detail.	82
Figure 4.3.1. Written example of “Vitalistic” Response Alternative Conception code.	88
Figure 4.3.2. Drawn example of Capillaries Absent Alternative Conception code.....	89

LIST OF ABBREVIATIONS

STEM = Science, technology, engineering, and math

IPT = Information processing theory

SOI = Selection, Organization, Integration

KC = Key Concept

MT = Mass transport

CRI = Circulatory and respiratory systems are physically integrated

CCE = Countercurrent exchange

AC = Alternative Conception

COH = Coherence

DoV = Degree of Visualization

DD = Depth of Drawing

ES = Expert Score

CHAPTER 1

INTRODUCTION

In 2011, the American Association for the Advancement of Science (AAAS) released a report entitled *Vision and Change in Undergraduate Biology Education: A Call to Action*, which detailed five core concepts and six core competencies that all general education biology curricula should seek to address (American Association for the Advancement of Science, 2011). This report emerged from a collective realization by biology faculty and scientists alike that the undergraduate educational structure for biological sciences was not adequately teaching authentic science practices that students needed for their future careers, especially in light of recent technologies that are rapidly expanding scientists' understanding of phenomena across various subdisciplines. A large-scale national effort was undertaken to build a consensus framework regarding undergraduate general biology instruction, resulting in the core concepts (evolution; structure and function; information flow, exchange, and storage; pathways and transformations of energy and matter; systems) and core competencies (applying the process of science; using quantitative reasoning; using modeling and simulation; understanding interdisciplinary nature of science; communicating and collaborating between disciplines; understanding relationship between science and society) (AAAS, 2011). Both implicitly and explicitly, *Vision and Change* suggested that undergraduate life science instruction needed to shift from rote memorization of lecture-based content with teacher-as-authority models, to more active, inquiry-based learning frameworks in which students are engaged in authentic science practices that develop and assess more critical thinking and problem-solving skills.

Because of the cross-disciplinary nature of marine science courses, they offer unique opportunities to integrate these core concepts and competencies within a singular instructional

framework at the undergraduate level. Numerous organizations and studies over several decades of research have advocated for the benefits of including marine science instruction in K-16 education (Brody & Koch, 1990; Cudaback, 2006; Cummins & Snively, 2000; Fortner & Mayer, 1989; Gough, 2017; Guest, Lotze, & Wallace, 2015; Lambert, 2005, 2006; Lucrezi, Milanese, Danovaro, & Cerrano, 2018; Strang, Decharon, & Schoedinger, 2007). Three major themes have emerged from this body of literature: (1) that marine science courses meet national instructional standards and benchmarks for science instruction; (2) that the integrative nature of marine science fosters deeper student understanding of cross-disciplinary topics; and (3) that there exists a high level of student motivation and positive societal outcomes when marine science courses are included in curriculum development. Despite these clearly articulated benefits, specific research into marine science education- whether regarding grade level, sub-discipline, curriculum development, or teaching strategies- has lagged significantly behind education research in other science, technology, engineering, and math (STEM) disciplines such as physics, mathematics, and general biology.

A marine science curriculum is an inherently integrative, cross-disciplinary approach to science education, as a single concept (such as ocean acidification) draws from and connects several subdisciplines including biology, chemistry, physics, and geology. Because of this cross-talk between content areas, marine science curricula provide opportunities to deepen student understanding of key science concepts, as well as promote higher-level critical thinking and problem-solving skills that directly address the call to action of the *Vision and Change* framework (Cummins & Snively, 2000; Gough, 2017; Lambert, 2005, 2006; Strang et al., 2007). Instructional strategies that have also led to more meaningful learning of science concepts, such as drawing-to-learn interventions (Heideman, Flores, Sevier, & Trouton, 2017; Quillin &

Thomas, 2015; Scheiter, Schleinschok, & Ainsworth, 2017; Wu & Rau, 2019), could therefore complement and enhance student learning in marine science. Learner-generated drawing activities could prove to be a useful strategy for students to tackle a variety of concepts within marine science, from drawing pictures to understanding anatomical differences between marine organisms, to primary productivity concept mapping, to diagramming physical features of oceans, such as water masses and currents. Furthermore, learner-generated drawings could bridge a crucial gap between instruction and practice by providing explicit opportunities for students to develop their visual literacy and science communication skills, both critical competency objectives if students are to authentically engage in science. To date, however, there appear to be no formalized assessments or studies applying such drawing-to-learn approaches to marine science education.

The goal of this study was to introduce drawing-to-learn strategies in a formative assessment setting to an introductory-level undergraduate marine science course at a large public university in the Northeastern United States, and to evaluate if such strategies are useful for both uncovering student knowledge on important biological concepts, and revealing how students communicate that knowledge. Integrating data collected from previous years in which no drawing activities were prompted or used provided a basis of comparison to help understand whether and how drawing-to-learn activities impacted student representation and communication of their knowledge of these concepts. This introductory chapter reviews marine science education and drawing-to-learn literature. The theoretical cognitive framework within which these concepts reside provides the context for examples of drawing-to-learn interventions used in biology and anatomy coursework across grade levels. Finally, an explanation of how and why

these generative learning activities align with the goals of *Vision and Change* and ways in which they can be implemented sets the stage for the current study design and execution.

1.1 The need for marine science in science curricula

Advocacy for including marine science concepts in standard K-12 science curricula is not new; organizations such as Sea Grant have recognized and prioritized funding efforts for teacher professional development and marine education curriculum development since the 1970s (Fortner & Mayer, 1989). Additionally, the United Nations' Sustainable Development Goals have included both marine conservation and education targets since 2015 (United Nations, 2018), and World Oceans Day has been officially recognized since 2008, and unofficially celebrated since 1992 (United Nations, 2019). In the United States, groups such as the Ocean Literacy Network have made strides to provide educators with resources to fit marine science education within existing curricular frameworks, such as their work to correlate ocean literacy topics with Next Generation Science Standards (NGSS, 2013) in a comprehensive matrix (Ocean Literacy Network, 2015; Strang et al., 2007). This matrix, and the corresponding report set forth by the National Marine Educators Association (NMEA, 2010), clearly detail ways in which marine science can be effectively integrated into standard science curricula. Indeed, incorporating marine science courses in high school curricula has been demonstrated to support established science literacy goals, yet the benefits of such courses are not recognized nationally (Lambert, 2006), nor internationally (Gough, 2017). The lack of progress in education research at the K-12 level is mirrored at the undergraduate level, with few reports available that assess post-secondary marine science education. As a result, despite efforts to underscore the utility and importance of the inclusion of marine science as a worthy educational goal for decades,

developing marine science curricula across all levels (K-16) is still often perceived as a grassroots effort in its infancy.

1.2 Marine science, meaningful science learning, and instruction

Due to their inherently integrative nature, marine science courses reflect the real-world environment of interdisciplinary science with instructional approaches that combine concepts from diverse fields as biology, geology, chemistry, and physics (Lambert, 2005, 2006). In Lambert's studies (2005, 2006) of high school marine science courses in Florida, students who received instruction from teachers who had deep understanding of marine science and followed a well-integrated instructional approach demonstrated deeper content understanding and better critical thinking skills than students in other programs. These studies highlight the need for more pre- and in-service training opportunities for teachers to help implement marine science education at the K-12 level, but also demonstrate the utility and efficacy of a well-executed marine science curriculum.

In another study of fourth-grade students in British Columbia, the authors assessed not only student knowledge, but also stances and values the students held before and after experiential, inquiry-based instruction in marine science (Cummins & Snively, 2000). Both content knowledge and positive attitudes towards the marine environment increased in students after these lessons, with students demonstrating greater understanding of the interconnected nature of the marine environment and real-world applications of their knowledge (Cummins & Snively, 2000). A separate study in Oregon assessed marine knowledge in 4th-, 8th-, and 11th-graders to understand whether and how students conceptually linked different ideas within subdisciplines across marine science in a meaningful way (Brody & Koch, 1990). The authors found a common theme across grade levels, in which students incorrectly connected concepts

across disciplines (e.g. linking a marine biological topic like fish breathing underwater incorrectly or only partially correctly to a chemistry topic like oxygen availability in the water) due to misconceptions (i.e., alternative conceptions) or preconceptions they held based on personal experience (Brody & Koch, 1990).

Beyond the variety of worthy educational goals that a marine science curriculum offers to meet, it is important to consider student motivation and the value of marine science as a field of study. A survey by Cudaback (2006) of students at North Carolina State University- Raleigh indicated that students had high motivation and interest in learning more about the ocean. However, most of the students' current knowledge and passions had been influenced by personal experience, anecdotal information from peers, and informal educational settings, thereby opening up more possibilities for misinformation and misunderstanding (Cudaback, 2006). Similarly, a study of 7th- through 12th-grade students in Nova Scotia showed that students personally valued the ocean and possessed a great interest in studying marine science topics, yet the average score on quizzes that assessed a variety of marine-related cross-disciplinary subjects was less than 50% (Guest et al., 2015). In Italy, university students studying marine science also indicated high motivation and excitement about the marine biology field, but had received very little by way of formal instruction prior to their enrollment at the university (Lucrezi et al., 2018).

Taken together, the results from these studies suggest that the interdisciplinary nature of marine science education can offer opportunities for students to engage in making deeper, cross-cutting connections in content areas that are typically siloed to individual fields. This integrative approach means marine science curricula are uniquely positioned to provide more meaningful learning opportunities for students in STEM. From a pedagogical standpoint, explicit instruction that directs student attention to the cross-disciplinary nature of many marine science concepts

may also more readily identify alternative conceptions or preconceptions that students hold, such as those identified in the Brody and Koch paper (1990). Simultaneously, students of all ages are motivated to study marine science, yet lack formal educational opportunities to do so prior to university. The marine science education literature overall lacks specific strategies and best practices for instructors to accomplish these goals, and a more robust research effort on effective teaching strategies is needed to help develop and inform marine science curriculum development, as well as better prepare pre-service teachers. By extension, a modification of science curricula to include more marine science instruction could therefore be more likely to attract and engage students with material than traditional programs that teach science subjects in isolation. Existing marine science courses and curricula, such as the one in this study, can also be examined and modified to promote deeper learning and address student preconceptions, as they provide a rich space for specific research-based instructional strategies to be tested and implemented.

1.3 The research deficit regarding marine science education

Despite all of the positive effects that inclusion of marine science courses has on both K-12 and university students' learning outcomes (e.g. Cudaback, 2006; Cummins & Snively, 2000; Lucrezi et al., 2018), much of the recent focus on marine science education has been with the aim of increasing ocean literacy to achieve environmental conservation and restoration goals (e.g. Gough, 2017; Guest et al., 2015; Ocean Literacy Network, 2013), rather than actual classroom practices that instructors can implement. Only a few studies were found thus far that directly assessed marine science instruction specifically and thoroughly across an undergraduate program (e.g. Arthurs, Hsia, & Schweinle, 2015; Barrett, Swick, & Smith Jr., 2014; Weatherbee & Lindsay, 2018).

A curriculum-wide assessment was undertaken at the United States Naval Academy (Barrett et al., 2014) to analyze the Oceanography program in the context of Hess's cognitive rigor matrix (Hess, 2006). This model allowed the authors to determine cognitive demand of each of the Oceanography classes on the 200-, 300-, and 400-level in order to see if courses were classified appropriately, as well as determine the level of rigor demanded of students by different instructors. The Arthurs et al. paper (2015) discussed the development and implementation of the Ocean Concept Inventory, a means by which instructors can assess student understanding of a variety of oceanographic topics that is both reliable customizable to a degree. Finally, the study by Weatherbee and Lindsay (2018) assessed student understanding of a single topic – primary productivity – across an entire marine science program in the School of Marine Sciences at the University of Maine, which provided instructors across course levels with critical insight into student misconceptions and highlighted gaps in instruction. These studies are valuable in their contributions towards better structuring and assessing marine science education at the undergraduate level, but unfortunately stand alone in the literature. Given the high self-reported student interest in marine science education (Cudaback, 2006), and the demonstrated instructional benefits of such curricula (Lambert, 2005), research that focuses on how best to implement and assess marine science instruction is critically needed at the undergraduate level.

1.4 Drawing-to-learn in the marine science classroom

Strategies that are successful in other STEM disciplines, particularly drawing-to-learn tasks and improving visual literacy, could prove to be effective ways to enhance marine science instruction and deepen student understanding. With the science communication core competency from *Vision and Change* (AAAS, 2011) in mind, drawing-to-learn may provide a means of achieving that end in marine science, especially considering the visual nature of many marine

science topics that require expertise in the field to understand and communicate ideas in diagrammatic form (see Figure 1.4.1 for examples). One content area within marine science that

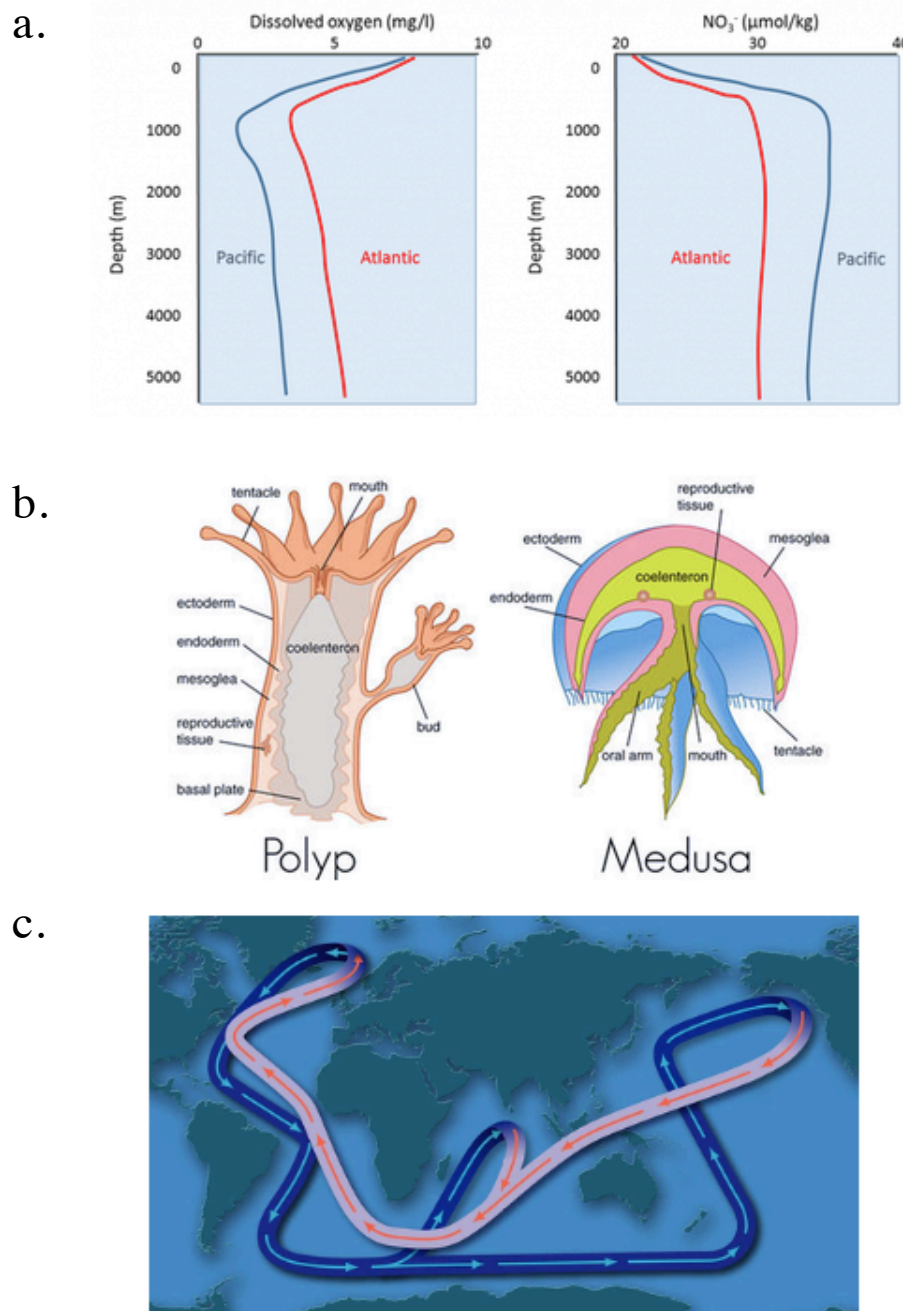


Figure 1.4.1 Examples of common diagrams found in marine science instruction and publications. Much of marine science content is visual in nature and requires explicit instruction in understanding key diagrams, such as those for (a) oceanographic depth profiles (Webb, 2020), (b) tissue layers in different stages of cnidarians (e.g. jellies and corals) (Wilkin & Blanchette, 2019), and (c) global oceanic circulation patterns (Doucette, n.d.).

relies heavily on visual representations to communicate concepts is that of comparative anatomy between marine phyla. Beyond merely contrasting different marine phyla and/or taxa and their biology, studying marine anatomy can provide important insights into larger-scale themes and ideas in biology. For example, a 2014 study on marine mammal tracheas offered insights into mammal dive depth and respiratory ventilation rates, but also provided evidence to support the *Vision and Change* (AAAS, 2011) core concept of structure following function (Moore et al., 2014). Likewise, a study of musculature and nervous system structures in brachiopods (a type of marine organism that resembles clams and mussels but belongs to its own phylum) suggested possible larger evolutionary implications based on observed anatomy (Santagata, 2011); evolution, too, is a core concept. Therefore, as one possible pedagogical approach for marine science, anchoring instruction and coursework in comparative marine anatomy can be a useful way to provide students opportunities to grapple with “big ideas” that are key to learning authentic skills and concepts. To do so, instructors need to ensure that students are equipped with the skills to interpret, understand, and use visual representations in their study of marine biology.

The purpose of this study was therefore to explore the utility of including drawing-to-learn activities in an undergraduate marine science course, in an effort to compare how students communicate their knowledge in written versus drawn form. This study focused on the following research questions:

1. How do students communicate their knowledge on a marine science topic when the question asks them to respond with a written versus drawn answer? (RQ1)
2. How do written versus drawn responses differ when evaluated on a novice-to-expert scale? (RQ2)

3. How does the nature of alternative conceptions or inaccuracies differ between response type? (RQ3)

Using coding methodologies to determine the level of quality of responses (as compared to expert-generated written or drawn responses), as well as levels of understanding (including inaccuracies or alternative conceptions), this study sought to articulate differences in student representations of their knowledge of circulatory and respiratory systems in marine organisms between written and drawn responses. With this data, the utility of including such drawing-to-learn strategies in an active-learning marine science classroom was assessed with the aim that such activities may be refined further for future implementation to better student understanding of important core concepts in marine science, as well as uncover student ideas and preconceptions that may arise.

1.5 Visual literacy and STEM education

1.5.1 Visual literacy

Before covering the specifics of drawing-to-learn strategies, which will be described in Section 1.6, it is important to understand how these strategies fit into the broader view of both visual literacy and science communication. “Visual literacy,” a term first coined in the 1960s, (Felten, 2010) is an important academic skill that has gained increasing attention in recent decades across different academic disciplines. As defined in his 2010 review paper on its foundations, Felten describes visual literacy as understanding, producing, and using culturally relevant, significant images in a way that is akin to literacy as it is traditionally understood in a text-based context. Felten’s review paper was published as the debate over how and why to include visual literacy as a worthy curricular goal was coming to a head; in 2012, the Association of College and Research Librarians (ACRL) published a comprehensive report that set forth

seven standards that define visual literacy competency in the context of higher education. Further, the report sets forth specific performance indicators that support each standard with specific learning outcomes. The ACRL refined the definition of visual literacy to be “the set of abilities that enables an individual to effectively find, interpret, evaluate, use, and create images and visual media” (ACRL, 2012, p. 97). The ACRL report and definition serve as the framework for visual literacy for this study.

1.5.2 Discursive fluency and addressing visual communication skills

The aforementioned science communication core competency described in the *Vision and Change* document (AAAS, 2011) includes not only written and spoken language, but also includes visual communication methods using representations such as models, diagrams, and graphs (Offerdahl, Arneson, & Byrne, 2017). In order to participate meaningfully in what Airey and Linder (2009) call “disciplinary discourse,” a student must become familiar and comfortable with using various “modes” of understanding the discipline, which can include spoken and written language, commonly used gestures, mathematical representations of phenomena, the tools themselves scientists use, and images, which could include graphs, diagrams, and/or pictures. This familiarity with modes of disciplinary understanding can be conceptualized as a “discursive fluency,” i.e. “the achievement of fluency in the disciplinary discourse scientists use when engaging in activities such as 1) decoding and interpreting visual representations, 2) encoding and creating visual representations, and 3) generating mental models” (Offerdahl et al., 2017, p. 2). If instructors are to heed the recommendations of *Vision and Change* (AAAS, 2011) and truly teach students authentic science practices, then courses and assessments should be designed to equip students with the skills and tools necessary to achieve this discursive fluency.

Although recognized as a critical part of scientific discourse, explicit instruction in visual literacy as a building block of discursive fluency is typically overlooked in undergraduate science education; oftentimes, science lecturers neglect to address or probe student understanding of representations because they themselves take for granted the fluency they already possess that is required to deeply understand and use visual representations (Airey & Linder, 2009). Furthermore, students are savvy about effort-to-reward ratios when it comes to learning and studying for their science courses; when instructors do not address visual literacy in assessments, be they formative or summative, it signals to students that it is not valued and thus they will not spend the time to learn these skills (Airey & Linder, 2009; Offerdahl et al., 2017). With this in mind, it is important for instructors to consider if adding active learning tasks as a part of routine coursework, such as drawing-to-learn prompts used in this study, could augment students' visual communication skills by giving them more practice generating and manipulating their own visual representations of knowledge.

1.5.3 Visual literacy in science

The use of images as models, data representations, and tools to visualize what is invisible to the eye is ubiquitous across science disciplines, yet science curricula rarely focus on in-depth, explicit instruction in the creation and interpretation of these images. In their 2015 paper, Evagorou, Erduran, and Mäntylä argue that visualization should be taught to science students with the intent to use visualization as a procedural strategy in itself for better understanding content. Students engaged in the *process* of creating images (rather than strictly interpreting the *product*) not only become better versed in using images to problem solve, build their own knowledge, and facilitate knowledge transfer, but also are given the opportunity to use and

manipulate visualizations that are used in real-world science professions and applications (Evagorou, Erduran, & Mäntylä, 2015).

People commonly – yet incorrectly – assume that living in an image-saturated world means that we possess inherent skills that qualify us as visually literate (Felten, 2010). This assumption is often on display in the science classroom, where teachers tend to skim over or simply point to parts of an image in science texts during instruction, without taking the time to explicitly teach students about the information the visual is attempting to convey, or about conventional diagrammatic tools used in different ways by different disciplines (e.g. arrows conveying the idea of force in a physics textbook, versus indicating a sequential process in a biology diagram) (McTigue & Flowers, 2011). In a study investigating student assessment and understanding of water cycle diagrams across textbooks that were targeting grades 2, 4, 6 and 8, results indicated that students often misinterpreted conventions, valued diagrams with lots of information (which they determined by the amount of words) but also a streamlined design (less clutter), and struggled to use the diagrams as anything but a concrete representation of the water cycle, even when prompted (McTigue & Flowers, 2011). These results underscore the importance of explicit instruction in visual literacy and the dangers in assuming that students understand images in STEM instruction simply because of high levels of exposure to them. In giving learners opportunities to practice and create their own diagrammatic representations of marine science content in response to a given prompt, this study seeks to advance the body of education literature to include specific strategies that may prove useful in other undergraduate classrooms in an effort to increase visual literacy amongst students.

1.6 Drawing-to-learn: strategies and instructional goals

1.6.1 Defining ‘drawing’ and ‘drawing-to-learn’

Drawing-to-learn strategies provide one avenue for instructors to train students in the skills they will need to achieve visual literacy and become more fluent in disciplinary discourse. The terms ‘drawing’ and ‘drawing-to-learn’ likely evoke different meanings for different readers; for instance, one person may consider a drawing to be constrained to a pencil-and-paper representation, whereas another person may consider a computer-generated image to be a legitimate drawing. Further complicating the understanding of these terms is the diversity of language used in the literature, with different authors using different terminology interchangeably with drawing(s) (e.g. sketches, diagram, visualization, illustration, etc.) (Quillin & Thomas, 2015). For the purposes of this study, a ‘drawing’ is described in terms of the definition set forth by Quillin and Thomas (2015): “a learner-generated external visual representation depicting any type of content, whether structure, relationship, or process, created in static two dimensions in any medium.” In considering the breadth of this definition, the authors point out four factors that should also be kept in mind: (1) a continuum exists of the degree to which a drawing is learner-generated, from starting with a blank piece of paper to interpreting a supplied diagram; (2) drawings are external representations of that which was initially formed internally (i.e. mentally; discussed in more detail below in Section 1.5); (3) drawings exist on a spectrum from faithfully representational to abstract; and (4) use of any two-dimensional medium qualifies as ‘drawing’ (Quillin & Thomas, 2015).

Much like the idea of drawing, ‘drawing-to-learn’ is a term that can be used flexibly depending on the instructional goal. In STEM, drawing-to-learn can be employed for a variety of educational goals that specifically develop skills “that are core to science thinking, including:

observation, problem-solving, explanation, and communication,” (Fan, 2015, p. 170). Ainsworth (2011) expands upon these categories by also including drawing-to-learn as a way to increase student engagement in STEM content, and more generally as a ‘learning strategy’ that engages cognitive strategies (see Section 1.8). Quillin and Thomas (2015) perhaps provide the most comprehensive breakdown of drawing-to-learn strategies, in which they categorize such activities in a matrix of pedagogical goals that are split between formative versus summative assessments across representational versus abstract modes of drawing.

In this study, drawings produced by students were a result of formative assessment questions asked throughout the semester at the end of lectures. Further, no explicit instruction was given as to whether drawings were to be representational or abstract- this was left to the discretion of each student, although it was emphasized that drawings would be assessed for content and not for artistic skill. The pedagogical goals of drawing-to-learn for this study spanned both the representational and abstract categories of formative assessments, including (but not limited to) fostering active learning, understanding of spatial relationships, construction of mental models, acquisition of content knowledge, and connection of concepts (Quillin & Thomas, 2015, Table 3).

1.6.2 Drawing-to-learn framework in STEM

In an evolving landscape of evidence-based strategies for planning and delivering quality science education, innovative learning strategies such as multimedia, image-based, and drawing-to-learn activities are gaining support as instructional activities alongside traditional skills such as reading and writing (Ainsworth, Prain, & Tytler, 2011; Fan, 2015; Quillin & Thomas, 2015; Scheiter et al., 2017; Wu & Rau, 2019). Such studies add to the research on image-based learning, which has included topics like support for students’ learning and utilizing a variety of

cognitive and metacognitive skills (e.g. Backhouse, Fitzpatrick, Hutchinson, Thandi, & Keenan, 2017; Edens & Potter, 2003; Fiorella & Mayer, 2016), demonstrated academic benefits of more visual learning tasks, including STEM-specific advantages (e.g. Gross, Wright, & Anderson, 2017; Mason, Pluchino, & Tornatora, 2013; Mautone & Mayer, 2007), and strategies for instructional design and implementation (e.g. Fiorella & Mayer, 2016; Leutner & Schmek, 2014; Mayer & Fiorella, 2014). Multiple studies have investigated how visual and generative learning strategies can enhance deeper understanding of science content (Edens & Potter, 2003; Gross et al., 2017; Mason et al., 2013).

For the undergraduate course from which data were obtained, several instructional goals were stated in the course syllabus, including students being able to “(1) recognize diverse marine organisms, their key characteristics and phylogenetic relationships; (2) understand the basic biology of marine organisms related to their functional adaptations to different habitats; and (3) be able to describe key processes in marine organisms...” (2019 course syllabus). Much like how the previous study examples (Moore et al., 2014; Santagata, 2011) went beyond the specifics of the research questions at hand to connect with larger themes within biology, these course goals aligned well with core concepts outlined in *Vision and Change*, notably “Evolution” and “Structure and Function” (AAAS, 2011). Open-notebook formative assessment prompts spanned a variety of topics and ideas throughout the semester, but several specifically focused on different key aspects of marine organisms and their anatomical differences that help distinguish between phyla, which targeted the specific learning outcomes outlined above. Since drawing has been shown to be especially useful for understanding certain biology topics (including anatomy) that lend themselves to visual-spatial reasoning (Fan, 2015; Scheiter et al., 2017), it was introduced as a strategy in 2019 to questions that targeted comparative anatomy in an effort to

understand whether such generative, drawing-to-learn activities facilitated students in demonstrating more expert-like communication and greater depth of detail of the concepts than in past years when only written responses were recorded. In particular, a prompt regarding circulatory – respiratory system linkage was chosen as the focus of this study, due both to the highly visual nature of the subject matter (see Figure 1.6.1) as well as its alignment to course goals and overarching *Vision and Change* (AAAS, 2011) core concepts and competencies.

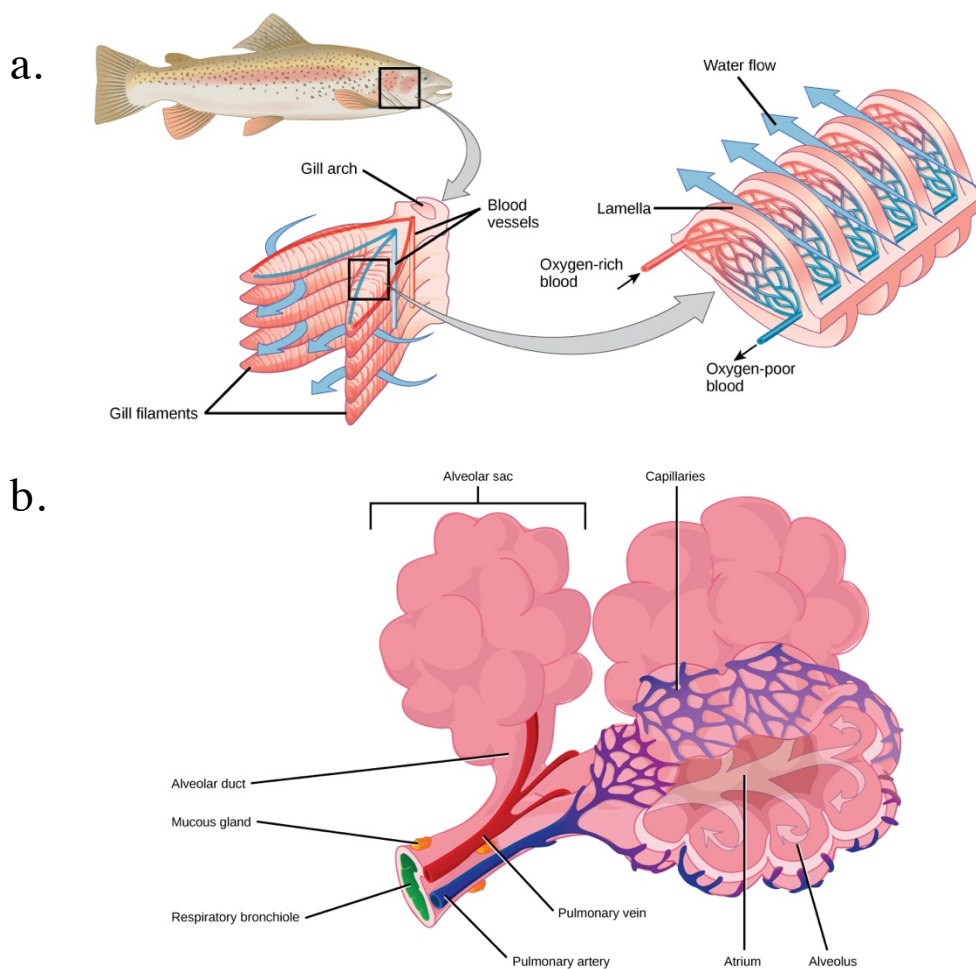


Figure 1.6.1 Examples of diagrams depicting the linkage between circulatory and respiratory systems. Marine science and general biology textbooks typically rely heavily on diagrams to represent conceptual information. Here, image (a) depicts the integration of capillaries in the gills of fish, and image (b) capillaries integrated in substructures (alveoli) of the lungs; both images from Molnar & Gair (2015).

With any classroom activity, it is important that instructors have a clear end goal and/or means of assessment in mind in order to effectively design instruction to equip students with the skills and conceptual knowledge needed to achieve that goal (Wiggins & McTighe, 2005). To that end, understanding the theoretical learning framework within which drawing-to-learn is situated, as well as the cognitive processes that drawing-to-learn supports and activates, can help instructors better design the means to achieve the ends. The following sections outline the information processing theory framework, with special attention to Mayer's cognitive theory of multimedia learning, to situate drawing-to-learn tasks in the wider educational discourse. Drawing-to-learn strategies activate and promote particular cognitive processes, and lend themselves to developing students' visual literacy; moreover, there are inherent STEM-specific benefits to implementing such strategies in the classroom.

1.7 Information Processing Theory

Information processing theory (IPT), an umbrella category that captures a multitude of different sub-theories, provides a useful lens through which to view and understand the utility of drawing-to-learn strategies. Information processing theories, though different from one another in how they describe the cognitive processes of information acquisition, translation, and storage, are rooted in several common assumptions that span each individual theory. IPTs all hold that cognitive processing is akin to a computer processing system: information is received, stored, and retrieved as necessary in a multi-stage process (Schunk, 2012). Further, learning is viewed as the encoding of information into long-term memory, a process by which the learner relates incoming information to prior knowledge so as to generate a meaningful construct or schema within which the new knowledge fits (Schunk, 2012).

In the context of drawing-to-learn strategies and the cognitive benefits provided, specific IPTs such as Mayer's cognitive theory of multimedia learning (Mayer, 2014) comprehensively describe the relationship between verbal and visual information, and how learning is enhanced with the inclusion of images versus just text alone. Mayer's multimedia learning theory builds upon previous information processing theories such as Paivio's dual-coding theory (Clark & Paivio, 1991; Paivio, 1991) and Wittrock's generative theory (Fiorella & Mayer, 2016; Wittrock, 1989), and is particularly useful when approaching drawing-to-learn strategies and the cognitive processes that they activate. In addition, specific learning strategies have emerged from Mayer's work that offer suggestions on how to best implement and execute image-based or drawing interventions in the classroom to maximize learning (e.g. Leutner & Schmek, 2014; Mayer & Fiorella, 2014).

1.7.1 Dual-coding and generative learning theories

Dual-coding and generative learning theories provide the foundation for Mayer's integrative theory. Dual-coding refers to the learner forging connections between verbal (text-based) information, or a text 'code,' and visual (image-based) informational code (Clark & Paivio, 1991; Paivio, 1991); Paivio contended that the effort exerted by the learner to make these associations naturally forged a deeper understanding of the material than provided by text alone (Mason et al., 2013). Similarly, Wittrock's generative learning theory (Wittrock, 1989) contends that learning is a process that includes generation, motivation, attention, and memory. This theory not only ascribed an active role to the learner, in how he or she makes sense of new information and integrates that knowledge into his or her existing mental model(s), but was also the first time educational researchers were provided with a framework with which to assess instructional design and predict subsequent performance outcomes (Fiorella & Mayer, 2016).

Learner-generated drawings specifically address some of the principles that are fundamental to both dual-coding and generative theories, as students are actively involved in the translation and encoding of verbal to visual material and vice-versa (Edens & Potter, 2003; Fiorella & Mayer, 2016; Mayer, 2014).

Both of these theories emphasize the connections or associations that are formed during the learning process between verbal, non-visual information, and visual information. Wittrock's theory, and subsequent learning theories that have developed from it, assign learning to three distinct processes of selection, organization, and integration of outside information into prior mental schemas; generative visual learning strategies working within this framework can thereby produce a more nuanced and deeper understanding of material than text-based learning alone (Fiorella & Mayer, 2016). Therefore, presenting visual and verbal information during instruction on marine science topics such as comparative anatomy, paired with drawing-to-learn assessment activities such as the formative question prompts used in this study, can provide students with the opportunities to engage in more meaningful learning.

1.7.2 Mayer's cognitive theory of multimedia learning

Richard Mayer, one of the most prominent theorists and researchers in the image-based, multimedia instructional design world, outlined a cognitive theory of multimedia learning that evolved from the work of Paivio and Wittrock (Mayer, 2014). Mayer organized his theory based on three assumptions: (1) two channels exist through which humans process auditory or visual information, (2) each channel possesses a limited capacity within which to process information, and (3) that learning is an active process by which a coordinated set of cognitive processes is activated (Mayer, 2014). In considering a multimedia presentation of material (i.e. one that combines verbal and visual information, whether through written text and/or verbal narration

paired with static images and/or animations), Mayer described five cognitive processes that the learner employs during such a presentation: selecting relevant words, selecting relevant images, organizing the selected words into a verbal mental model, organizing the selected images into a pictorial model, and finally, integrating these models with one another and into a mental framework that includes prior knowledge (Mayer, 2014). These processes, which can be summarized as Selection – Organization – Integration (SOI; discussed in detail below) provide the framework for a variety of studies investigating the utility of visual or image-based learning (Ainsworth et al., 2011; Edens & Potter, 2003; Leutner & Schmek, 2014; McCrudden & Rapp, 2017). Mayer’s cognitive theory of multimedia learning provides a lens through which image-based learning can be viewed, particularly in relation to the specific cognitive processes, including SOI and metacognition, that are activated by various visual learning strategies (e.g. Backhouse et al., 2017; Leutner & Schmek, 2014; McCrudden & Rapp, 2017).

1.8 Cognitive processes activated by imaged-based learning

1.8.1 Selection – Organization – Integration

The Selection – Organization – Integration (SOI) process is useful to consider as the backdrop for how students received information in multimedia format prior to answering the questions prompt in this study, as well as how their actual responses were communicated. Studies using drawing interventions in elementary classrooms have found higher learner performance (as measured by recall and inclusion of relevant information) in groups using generative drawing strategies versus those who either only read the text, or use written explanations (Edens & Potter, 2003; Leutner & Schmek, 2014). Ainsworth et al. (2011) hypothesized that the higher learner performance is due to the drawing process increasing learner motivation and engagement with the material. In other words, students who draw as a means of

understanding text-based information are engaging in a process of selecting relevant information from the text, creating a verbal mental model that is then translated into a pictorial mental model (which involves organizing and integrating the information into prior existing mental frameworks), and then translating the pictorial mental model into an external drawing (Leutner & Schmek, 2014). Thus, educators can strategically design classroom activities to encourage SOI cognitive processes, either implicitly or explicitly, to deepen student understanding of the material. This may prove especially useful for science text, which tends to be expository in nature and therefore more difficult than narrative texts for students to mentally access and process by reading alone.

Both Edens and Potter (2003) and Leutner and Schmek (2014) demonstrated higher student performance after drawing to learn versus simply re-reading text or taking notes; several hypotheses may explain these results. One such hypothesis posits that because sketching is a constructive, generative task, learners who use drawing techniques to help understand the information conveyed in text are inherently employing Mayer's Selection – Organization – Integration methodology when choosing what to include in their drawings (Scheiter et al., 2017). Another hypothesis as to why drawing leads to a richer understanding of material is that a pictorial representation of information may inherently be a better representation of some concepts, particularly those that are more spatial-visual in nature, than simply a text description alone (Fan, 2015; Scheiter et al., 2017).

To explore these two different hypotheses, Scheiter et al. (2017) prompted students to use either a verbal self-explanation strategy or a generative sketching strategy when reading an expository science text passage. Though, when taken as a whole, drawers did not outperform self-explainers on recall tasks, there was a difference seen between groups when assessing only

those responses from each group that were classified as high-quality (Scheiter et al., 2017). These results suggest that even those students who excelled in performance in the self-explanation group were still out-performed by the top-level students in the drawing strategy group, lending support to the second of the two hypotheses (experimental design flaws prohibited the authors from drawing conclusions about the first) (Scheiter et al., 2017).

With these studies in mind, it is important to note that lectures delivered prior to the circulatory – respiratory question prompt contained a blend of visual information (including diagrams, photos, and flow charts) as well as verbal information (text on the slides themselves as well as the lecturer’s narration). Students needed to distill several lectures’ worth of material to formulate articulate and robust responses, and oftentimes complex ideas that were better captured visually were included in model example diagrams on slides to facilitate student learning and expression of these concepts. Given these conditions, SOI is particularly relevant in considering the research questions at hand and the context within which student responses were generated. Although an examination of student notebooks as a means to assess what students actually wrote down from the lecture information was beyond the scope of this study, students had ample opportunity to engage in the SOI process, both while taking notes from the multimedia lecture, as well as while responding to the question prompt with the use of their notebooks. Ultimately, what students chose to communicate on their responses – both visually and verbally – is a partial reflection of the means by which students determined which information was relevant to include (i.e. the selection process), how students organized discrete pieces of conceptual knowledge spanning multiple days of lecture, and how information was integrated to form a response. Additionally, the open-notes format of the formative assessment probe allowed for students to potentially focus on deeper ideas and connecting concepts, rather

than a focus on surface-level factual recall (e.g. Eilertsen & Valdermo, 2000). In contrasting written responses to those that are drawn in light of the SOI process, this study sought to determine if more nuanced, deeper answers were obtained in one response format over the other, and to what degree such answers were verbalized versus visualized to assess the utility of prompting students to draw for this particular prompt.

1.8.2 Metacognitive skills

Metacognitive skills are also important to consider when analyzing drawing-to-learn strategies. Metacognition refers to the process of ‘thinking about thinking,’ or a conscious awareness of cognitive processes that can impact learning (Schunk, 2012). In particular, when a student is deciding which pieces of information from the text are important to include in a drawing, and how best to represent that information, he or she is inherently engaged in a metacognitive process of monitoring, reflecting upon, and editing the drawing (Leutner & Schmek, 2014). Explicit instruction in such metacognitive skills when introducing drawing strategies in the classroom can prove especially useful. When implementing a drawing technique dubbed ORDER (Observe, Reflect, Draw, Edit, Repeat) to a first-year medical school anatomy class, one study found that giving explicit instruction to reflect on their drawing, emphasizing learning outcomes instead of artistic output, and allowing students open-ended time frames in which to complete the task promoted metacognition and potentially deeper learning of the subject (Backhouse et al., 2017). Though this study in particular was seeking methods to reduce or replace costly dissection activities, there are still valuable instructional design considerations and metacognitive benefits to be gained from the results.

Finally, drawing-to-learn strategies can be particularly useful learning tasks since they require students to explicitly articulate their understanding of material in a visual way that can be

observed and critiqued by others, instead of ‘invisible’ mental models that may contain misconceptions or inaccuracies. For science disciplines in particular, drawing-to-learn activities that are well-planned and well-executed allow students to engage with the material on a deeper level as they distill textual information to the critical key points to include in their drawings (Ainsworth et al., 2011). In another study using a particular drawing technique called minute-sketches with folded lists (MSFL), the authors found that students using the MSFL strategy had a 20-50% higher recall score than students using strict visual review (VR); importantly, the MSFL strategy contained a few different checkpoints at which students reviewed both their drawings and listed keywords and checked for inaccuracies (Heideman et al., 2017), akin to the open-notes format of the question prompts in this study. In designing and implementing effective drawing-to-learn strategies, it is critical to include such feedback or collaboration steps, be it through self-reflection, instructor feedback, and/or comparisons to expert drawings or reference materials so students can identify such inaccuracies or misconceptions in their own knowledge (Fan, 2015). In this study, the aim was to assess if the metacognitive benefits of drawing-to-learn, compounded with similar benefits from open-notes assessment formats that allow students to focus on the *how* and *why* of a question instead of just the *what* (Eilertsen & Valdermo, 2000). potentially revealed themselves in the quality of drawn responses versus written.

1.8.3 Visual literacy and cognitive theory

In light of these important cognitive processes activated by drawing-to-learn, it is helpful to revisit visual literacy and how these ideas tie together. Visual literacy as defined by the ACRL report (2012) can essentially be summarized as the skills one needs to be able to effectively decode and critically think about the information that is presented in an image. Since a visually literate person taps into a set of cognitive tools to help enable this decoding, viewing visual

literacy through the lens of the cognitive theories discussed above may be useful for instructors (Beatty, 2013). For example, consider Standard Three from the ACRL report: “The visually literate student interprets and analyzes the meanings of images and visual media” (ACRL, 2012, p. 100). Dual-coding theory can be applied to the first performance indicator (“The visually literate student identifies information relevant to an image’s meaning” (ACRL, 2012, p. 100)), whereas it may be useful to demonstrate to students how the multimedia learning theory can serve as a guide for the second performance indicator (“The visually literate student situates an image in its cultural, social, and historical contexts” (ACRL, 2012, p. 101)) (Beatty, 2013). When assessing student responses, these links between visual literacy and the cognitive processes involved served as a useful lens for determining the quality and expertise of answers against the backdrop of information presented in lectures.

1.9 Implementing meaningful learning strategies: visual literacy and drawing-to-learn

Graphic representations, be they diagrams, data graphs, or illustrations, are an important part of communication and understanding across STEM fields (Evagorou et al., 2015; McCrudden & Rapp, 2017; McTigue & Flowers, 2011). With biology in particular, the temporal and spatial nature of many concepts lend themselves to visual representations that may enhance understanding more so than simple verbal explanations or text can (Fan, 2015). Additionally, effective science communication can be facilitated by well-designed visual displays that increase learner understanding and correct misconceptions (McCrudden & Rapp, 2017). Integrating drawing techniques in the life science classroom can develop students’ visual literacy skills, as well as provide learners with novel ways to approach and understand science content by using graphic representations to foster deeper learning, predictive skills, and problem-solving

techniques (Ainsworth et al., 2011; Edens & Potter, 2003; Heideman et al., 2017; Quillin & Thomas, 2015).

In considering all of the benefits demonstrated by drawing-to-learn strategies discussed above, many have argued that drawing should be considered an essential skill to be taught alongside more traditional skills (like reading and writing) in science curricula (Ainsworth et al., 2011; Fan, 2015; Quillin & Thomas, 2015; Wu & Rau, 2019). When introducing drawing-to-learn activities, however, it is important to consider the educational goals, as not every discipline or topic lends itself to drawing strategies (Ainsworth et al., 2011). Biology, for example, tends to lend itself well to visual interpretations that may be better suited to convey important information that text alone cannot sufficiently portray (Fan, 2015). As a whole, however, several learning processes have been identified in the cognitive and sociocultural research literature that are stimulated by drawing tasks and have been categorized as the following six types by Wu and Rau (2019): generative learning, self-regulation, mental model integration, spatial cognition, mediated discourse, and disciplinary practices. These skills, the authors argue, have been demonstrated time and again to foster a deeper understanding of STEM content for students and professionals alike, lending support for the integration of more drawing in STEM education (Wu & Rau, 2019). Finally, drawing-to-learn strategies and engaging students in visual learning can promote graphic literacy, which is important in both the interpretation and communication of data that is the cornerstone of much scientific discourse (Fan, 2015; Quillin & Thomas, 2015). For all of these reasons, given the visual-spatial nature of the question prompt in this study, the overarching goals of *Vision and Change* (AAAS, 2011), and the instructional goals within the marine science course itself, introducing drawing as a strategy for students to communicate their knowledge on the topic was seen as a worthwhile endeavor.

1.9.1 Diversifying assessment: the value of question type

Equally important to consider when deciding to integrate more visual literacy strategies into the classroom is how to ask good questions and structure assessments that align with desired learning outcomes. Considering the impetus for the *Vision and Change* document (AAAS, 2011) once more, much assessment focus in science courses in the past has been misaligned with stated learning outcomes – which may indicate the instructors’ desires for more critical thinking and novel problem-solving – instead mirroring how those instructors were themselves taught with a detail-oriented “facts first” framing of science content (Momsen et al., 2013). Furthermore, interview data from students indicated that they too desired to see more open-ended assessment questions in coursework, rather than the usual finely-focused multiple-choice formats that often require a mere recitation of rote-learned facts (AAAS, 2011). If crafted carefully, open-ended, constructed-response question types can provide a more authentic representation of student knowledge and understanding, as they tend to give students more opportunity to elaborate on their thinking and make their cognitive processes “visible” (Goubeaud, 2010).

Although constructed-response questions are not as easy or expeditious to grade in large lecture classrooms as are multiple choice, alternative question type options exist somewhere in between the two that still allow instructors to gain insight into student thinking. Research into a question format known as Multiple True-False (MTF) has revealed that careful wording of question stems can uncover student thinking and possible alternative conceptions they hold (Couch, Hubbard, & Brassil, 2018; Hubbard, Potts, & Couch, 2017). MTF questions have a similar format to multiple choice in that they have a question stem followed by multiple options, but students are asked to assess all of the answer options and determine if the statement presented is true or false for each. In addition to helping identify alternative conceptions that may

be amongst the answer choices like a free-response question could, MTF question types are faster to grade and may be more appealing to university faculty teaching large class sizes (Couch et al., 2018).

Beyond just question type for a singular assessment, it is important for instructors to consider administering multiple, varied assessment questions that speak to multiple facets of understanding a concept in order to gain as complete a picture as possible of student understanding (Schönborn & Anderson, 2008). Although an open-ended question type can provide instructors with much more information about student conceptions (whether accurate or inaccurate), no single question or question type can give a student the opportunity to demonstrate the full scope of his or her content knowledge and understanding (Schönborn & Anderson, 2008). Ultimately, the goal for instructors should be to have students' conceptions and understandings be as closely aligned to those of experts in the field; teachers should reflect on what knowledge and skills are required for "expert thinking" about a concept and design instruction and assessment backwards from there to ensure varied question types (both in structure and content) that cover these aspects (Wiggins & McTighe, 2005). A useful tool employed in this study that can help accomplish this is comparing student responses to exemplar expert responses, which creates a rubric or roadmap for instructors to use in determining where students may fall short in their responses. Teaching explicitly to these shortcomings can help address misunderstandings or misconceptions and align student thinking and reasoning to more expert-like discourse, thereby better preparing students for postgraduate work in STEM fields.

This study therefore sought to bridge the gap between the preponderance of evidence that has been discussed thus far about the meaningful learning advantages of drawing-to-learn tasks, as well as the importance of implementing more open-ended assessment questions to uncover

more telling insights into student knowledge in a marine science setting. Targeting a specific concept that is central to comparative anatomy in the marine biology discipline – that of circulatory and respiratory system linkages – the research questions presented here aimed to dig deeper and refine already-established constructed response question prompts in an existing marine science course by introducing the potentially useful and applicable tool of drawing-to-learn. By carefully considering which question prompts lent themselves to being diagrammed or drawn, and prompting students *explicitly* to communicate their ideas in a visual representation, this study aimed to provide a means by which marine science instructors could consider retooling their own assessment questions to more explicitly address visual literacy, science communication, and larger core ideas within the discipline.

CHAPTER 2

METHODS

2.1 Data collection

Data were collected in the form of individual formative assessment lecture “notecard questions” from students participating in an introductory marine science course at a public university in the Northeastern United States during the spring semesters from 2017 to 2019. The particular course served as a prerequisite for many students in the marine science major, and first- and second-year students were the majority of those enrolled each semester. For all years in which data were collected, students were informed at the beginning of the semester that research was being conducted, and of their option to opt out of having any of their work included for research purposes. Because the lecture notecards under study were completed as part of class participation credit, all notecards collected from each year were included in the study unless a student opted to have their work excluded, or a student was under 18 years of age at the time of data collection. No penalties were incurred for students who opted out of having their data included; however, completion of notecards was expected for all students as a part of course expectations outlined in the syllabus. Because only de-identified responses were analyzed and the key was destroyed, the subsequent research was deemed not to involve human subjects, and further Institutional Review Board approval was not required. Responses to notecard question prompts were collected periodically throughout the semester at the end of the 50-minute lecture. These questions were used as a means of checking student understanding on key concepts that had been covered in that day’s lecture as standalone topics, or those that built upon the previous two to three class periods’ worth of course content. Students were given roughly five minutes at the end of class to respond to the prompted question on an index card they received at the

beginning of class, and were able to look through their notes if needed as they responded. In 2017 and 2018, lecture notecard questions asked: "Describe an example to support the following statement: Respiratory and circulatory systems are closely linked." For 2019, the question was changed in order to explicitly prompt drawn or diagrammed responses, as follows: "Respiratory and circulatory systems are closely linked. Draw a diagram or sketch that shows this and label or describe it."

Between 2017 and 2018 ("written" years), 137 total notecard responses were collected; 2019 ("drawn" year) had a total of 79 responses. Student responses were collected, scanned, and de-identified, with each notecard receiving a code that corresponded to the year, topic of notecard, and a randomly generated three-digit student identification number (e.g. a 2019 notecard for the Respiratory and Circulation question for student 104 would be named "19R&C_104.pdf"). Identification keys were kept on a secure Google drive and were destroyed following the study. Only de-identified data was uploaded to Dedoose, a secure web-hosted qualitative analysis coding software (www.dedoose.com), and only the primary investigator had access to the project on the Dedoose platform. Students who withdrew from the class in previous years of the study but re-enrolled in later years were given unique ID numbers for each year but also had their previous year(s)' ID number noted on the lecture notecard to be able to link responses as necessary for analysis. Students who withdrew from the class had their ID numbers in blue typeface, whereas all other students had red typeface.

2.2 Grounded theory and multimedia discourse analysis

Analysis of notecard responses was situated within a blend of grounded theory and multimedia discourse analysis frameworks. Grounded theory was understood to be aligned with Strauss's conceptual framework, as described by Corbin (2013), in which *in-vivo* codes were

generated based on student responses, and refined/collapsed into larger categories throughout the analysis process. Multimedia discourse analysis was understood as described by Lemke (2015), in which a multi-modal “language” of images and text from student responses was analyzed for meaning, with the context of the class instruction, setting, and notecard activity itself considered during analysis. The blend of the two research methodologies allowed for a coding framework to be constructed such that student artifacts were considered as presented (grounded theory), but also with visual and textual meaning-making considered within the context of the course itself (discourse analysis), which provided a more comprehensive description of student knowledge – both content and quality – from the data than either methodology alone (e.g. Johnson, 2014). Image-based coding also required contextualization of diagrams within course material presented, as well as the coder’s own knowledge of the science content, and was based on the steps outlined in Mey and Dietrich’s Visual Grounded Theory Methodology (VGTM) (2017). Coding was therefore a highly iterative process, born out of close inspection and consideration of the data at hand; codes were generated in anticipation of the data, during data analysis, and even after an initial draft of data analysis, in which new codes needed to be generated to help fully tease out trends and comparisons in the data. The codebook that was ultimately constructed was an exploratory approach to this particular subject matter within marine science, as no similar approaches were found in the marine science education literature.

2.3 Determining expert versus novice responses

In assessing responses, “expert” responses were first constructed with guidance from the instructor to be able to compare student responses to the key concepts that the instructor was seeking to elicit in their answers. Although no singular “correct” answer exists for the question

prompts, a sampling of written and drawn responses were constructed ahead of time that capture possible expert answers. Examples of these written expert responses can be found in Table 2.3.1.

“Describe an example to support the following statement: Respiratory and circulatory systems are closely linked” (2017 & 2018)

Possible expert responses

In fish, lamellae of the gills are filled with capillaries of the circulatory system, and blood flows through these in opposite direction to water flowing over the gill filaments, in order to maximize exchange of gases.

In mammals, the alveoli of the lungs are surrounded by networks of capillaries from the pulmonary circulation so oxygen and carbon dioxide can be readily exchanged.

In birds, air capillaries from the parabronchi are interspersed with capillaries of the circulatory system to promote gas exchange.

Table 2.3.1. Possible expert written responses. Examples of written responses to the 2017 and 2018 notecard question prompt that satisfy “expert” criteria.

To provide additional examples of what a drawn expert response could look like, a focus group of graduate students from a required 500-level general marine biology class were also asked to construct a drawn response to the 2019 question prompt (see Figure 2.3.1). These expert responses were considered to be such as they aligned with Airey and Linder’s conceptualization of expert thinking (2009): high interconnectivity between ideas, thinking and reasoning across biological scale (from “invisible” processes such as gas exchange to whole-organism considerations), and contextualization of processes such as countercurrent exchange.

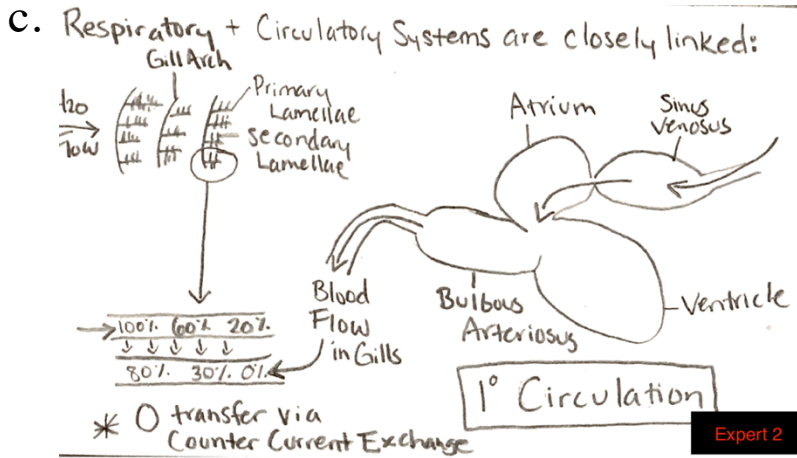
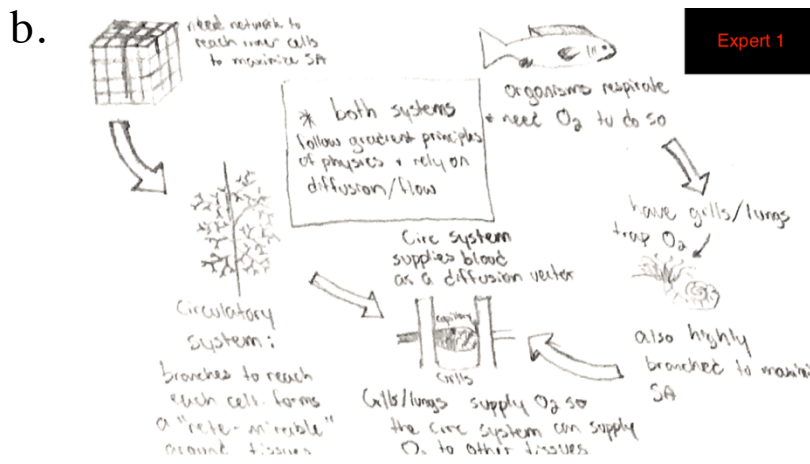
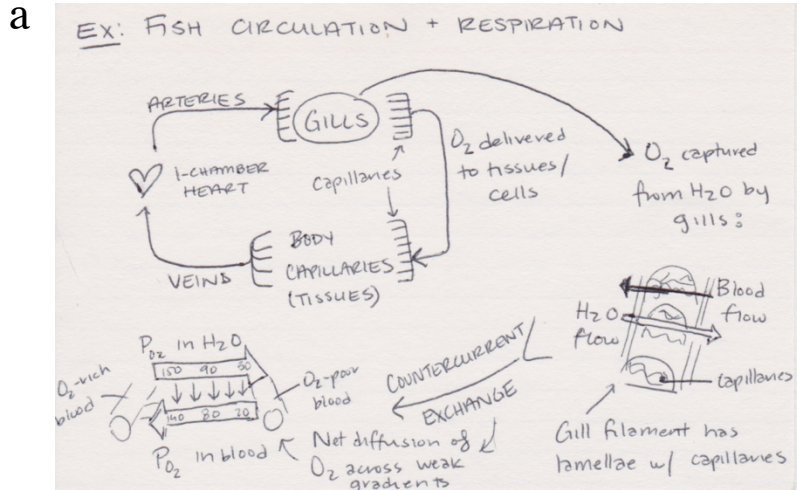


Figure 2.3.1. Possible expert responses for 2019 drawn prompt. Expert responses constructed by C. Siddons and S. Lindsay (a), and two graduate students enrolled in a 500-level graduate marine biology course “Expert 1” (b), and “Expert 2” (c). Expert responses reflect a systems-thinking approach with ideas highly interconnected; close-up details such as capillaries invested in gills and/or countercurrent exchange are also contextualized within the larger organism.

2.4 Codebook construction

To assess notecards for content and complexity of knowledge communicated, a codebook was drafted by C. Siddons and edited by S. Lindsay, following guidelines and a similar coding structure used by Dr. Asli Sezen-Barrie in her research (Sezen-Barrie, Stapleton, & Miller-Rushing, 2020; see also Appendix A). Although the subject matter between the two studies differs, understanding and approaching coding information on images (as opposed to text-based coding used in analyzing interview data, for example) was relevant and applicable. Coding was completed incrementally to assess inter-coder reliability as described by Saldaña (2016), first by applying the coding scheme to 10 random notecards. Two coders (S. Lindsay and C. Siddons) assigned codes separately, compared their results, and discussed any differences of interpretation behind the meaning of codes to assess reliability as well as refine the codebook language. Next, 20 notecards were coded and compared in a similar manner; by the end of this second round, codes assigned by both coders had reached 100% agreement. Thereafter, the remaining notecards were coded entirely by C. Siddons since the coding framework was comfortably established for reliable and consistent coding between notecards.

The three research questions were considered when constructing the codebook and what it should include as major categories (i.e. parent codes). All of these codes are discussed in detail in the following sections, as well as outlined extensively in Appendix A, but rationales for why codes were chosen in respect to each of the research questions are presented here. First, Research Question 1 (RQ1) was an effort to understand *how* students communicated their knowledge, and thus, needed a means to code for *what* students put on their notecard responses. To address this, the bulk of coding fell under the Key Concept (KC) and Representation parent codes. Key Concepts were determined to be the essential, “must-have” ideas that students should be

conveying in their responses to adequately demonstrate understanding of respiratory – circulatory linkage. Key Concepts were topics or ideas that were threaded through the discussion of not only the lecture content the day the notecard question was posed, but also in the preceding several lectures (roughly a week and a half of total class time) leading up to the notecard question, based on slide handouts from the 2019 course (S. Lindsay, personal communication). The Representation parent code was broken out into sub-categories child codes of structures, processes, and relationships, to better determine finer-grained detail in student responses that could accompany and/or support Key Concepts. Coding for and analyzing these two parent code categories allowed for drawing conclusions about what pieces of information students selected from lecture and determined was important in conveying their responses, particularly in light of the S-O-I framework (Mayer, 2014) discussed previously.

Research Question 2 (RQ2) required a means to assess where student answers fell on a spectrum from novice to expert. In order to facilitate this assessment, both Coherence and Expert Scores (ES) were applied to notecards at the end of all other coding. Coherence scores were awarded to give a sense of how interconnected student ideas on the notecard were, as indicated by grammatical and/or diagrammatic linkers between ideas in order to assess if there are any “gaps” between student thinking. The Expert Score provided a means of comparison all responses across years, as well as a standard against which cards were scored beyond the individual representations and pieces of knowledge students decided upon when constructing their responses. This score went beyond coherence to consider, given the requisite biology knowledge for the question itself, the level of interconnectedness between ideas presented on the notecard (e.g. how well students synthesized discrete pieces of knowledge presented over several class sessions) as well as how students represented knowledge across both scales and systems.

Finally, Research Question 3 (RQ3) sought to understand the nature of alternative conceptions between written and drawn responses, and if these were the same or different regardless of response type. Thus, a parent code of Alternative Conceptions (AC) was necessary to code for the different kinds of ACs present in student responses, which comprised the catalogue of child codes. These Alternative Conception child codes mostly arose from the data itself, rather than being constructed ahead of time, as it was difficult (if impossible) to predict what kinds of preconceptions or incorrect ideas students would convey in their responses. For fully coded examples of both written and drawn responses, refer to Figure 2.4.4 and Figure 2.4.5.

2.4.1 Key Concept codes

The first step in constructing the codebook was to determine what would be identified in student responses as Key Concept codes, or KCs. Mass transport (MT), circulatory and respiratory system are physically integrated (CRI), and countercurrent exchange (CCE) were determined to be the three KCs that were part of an expert response. Mass transport meant that either in writing or drawing form, students explicitly conveyed the understanding that circulatory systems serve to link gas exchange surfaces with cells in tissues over distances too great to cover by diffusion alone. In diagram form, this was understood to be conveyed if a student drew a circuit diagram linking the heart and/or other circulatory system features with the lungs/gills and/or the body (Figure 2.4.1a). For the second KC – circulatory and respiratory systems are physically integrated – students' responses needed to capture proximity relationships such that a clear understanding of the close association of capillaries with gas exchange surfaces was explicitly demonstrated (Figure 2.4.1c). In drawn responses, this was considered to be adequately conveyed if capillaries were obviously drawn or labeled in the lungs or gills.

Finally, countercurrent exchange was considered a bonus KC; that is, it was considered to be demonstrative of higher-order thinking and understanding of the linkage between the two systems (*if* it was contextualized and not presented as a standalone concept), as this type of exchange is more efficient than *concurrent* exchange due to diffusion of gases across weaker concentration gradients between blood and water. Student drawings were considered representative of countercurrent exchange if there was a clear indication of opposing water and blood flows, and concentration gradients (Figure 2.4.1b). Diagrams or descriptions of countercurrent exchange were coded for this KC whenever it was mentioned or depicted; however, other codes (such as Coherence and Expert Scores, described below) determined the level of contextualization of CCE and if having this “bonus” concept truly demonstrated expert-like thinking.

Each notecard was coded for one or more KCs if they were clearly described or shown as indicated above. Co-coded with each Key Concept code was a Use of Evidence code that indicated whether words, drawings, or both were used in the student’s response, as well as whether the student used only words when actually prompted to draw (2019 only; see Appendix A for full coding structure).

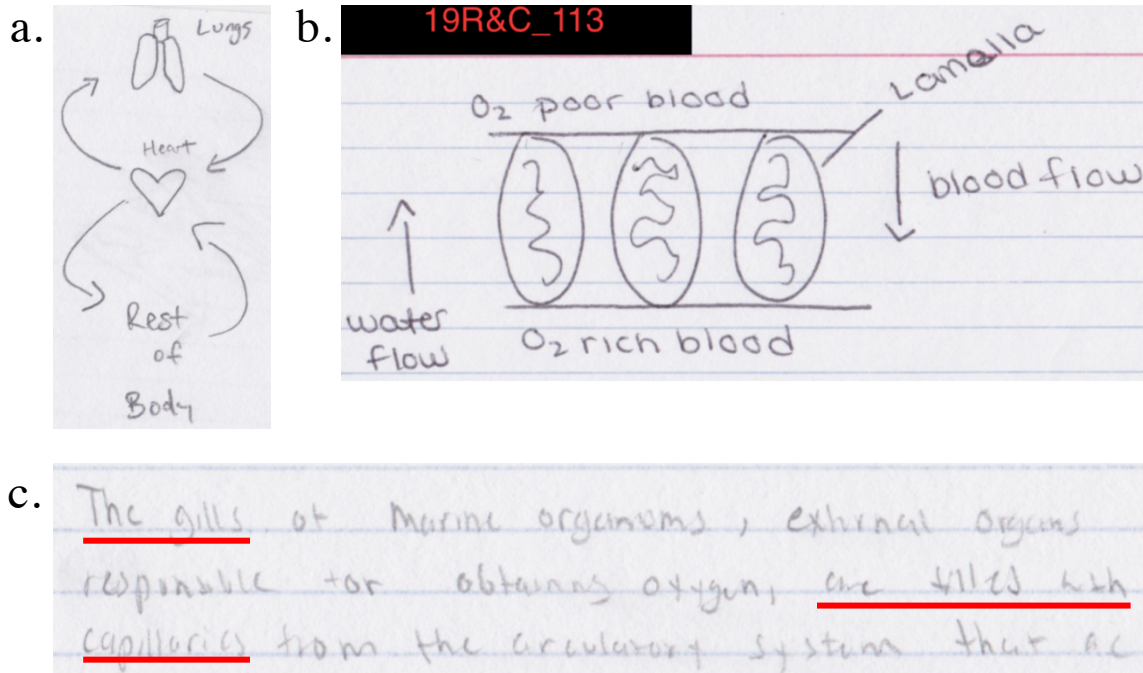


Figure 2.4.1 Key Concept student examples. Examples of student responses demonstrating (a) mass transport (drawn), (b) countercurrent exchange (drawn), and (c) respiratory – circulatory linkage (written; highlighted sections read: “The gills... are filled with capillaries...”).

2.4.2 Degree of Visualization and Depth of Drawing

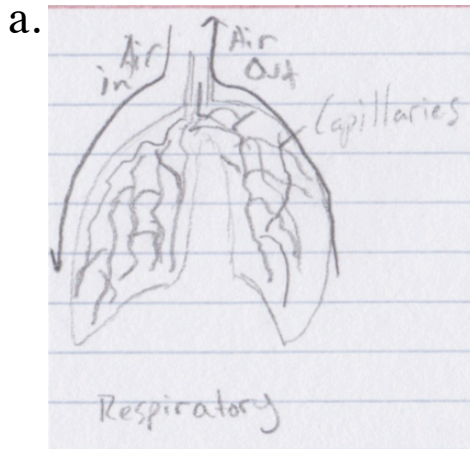
Every notecard was coded for two main parent codes that each had child codes of various levels: Degree of Visualization (DoV; 0 – 3) and Depth of Drawing (DD; 0 – 3 and N/A). Each of these scores was given a separate box in the top right or top left corner (depending on space) of the notecard. Degree of Visualization was also assessed on a whole-notecard approach, with a score of 0 indicating that only words were used to convey ideas; for years in which drawing was not prompted (2017 and 2018), students all received a DoV score of 0, which did not count negatively towards any assessment of coherence or accuracy of response. The DoV score was mostly used to get a sense of differences in visualization techniques and strategies used by students in 2019 in their drawn responses to potentially assess challenges students may have faced in attempting to depict their responses in a sketch. For example, low DoV scores in 2019

could indicate that students either included superfluous or minimal drawing efforts and/or that they continued to rely on written explanations of ideas despite being prompted to draw.

The Depth of Drawing (DD) score assessed the utility, detail, and accuracy of drawings used in a student's response, particularly in assessing the student's grasp of Key Concept codes. In considering what would be an 'expert response,' different levels of drawings were established that corresponded to these scores. "Baseline" drawings depicting circuits of material moving between the heart, lungs/gills, and/or body were considered on the lower end of the DD score, whereas efforts to capture the close integration of systems via capillaries in the lungs or gills, or contextualized depictions of countercurrent exchange, were given higher scores.

2.4.3 Representations

After Key Concept codes, Degree of Visualization, and Depth of Drawing were assessed, specific representations that students used to convey their knowledge were coded. Students who did not meet the threshold for depicting any of the three Key Concepts still had their responses coded for Representations, as it was possible for them to demonstrate some degree of knowledge about the two systems without accurately describing or depicting these larger ideas. The Representations category was comprised of a list of recurring elements seen in drawn responses, or explicitly named or described in the written responses, in order to get a sense of conventions, vocabulary, and/or mechanisms by which students were attempting to convey their understanding of the question posed. Coding for this data allowed for some degree of interpretation about what information students considered most important, and for drawn responses, what may have influenced their diagrams and drawings. Representations included



b. They are closely linked because the alveoli in the lungs are coated in capillaries & O_2 diffuses into the blood

Figure 2.4.2 Student examples of various Representations codes. Student artifacts that demonstrate several Representations child codes: Capillaries, Arrows (Directionality), and Lung/gills (**a**, drawn); Capillaries, Lung/gills, Written Molecule [O_2], and the process Gas Exchange (**b**, written).

physical aspects of the systems, such as the heart, lungs or gills, capillaries, and pulmonary and/or body circuits, clearly depicting or describing the process of gas exchange, as well as more abstract conventions such as using arrows to display a directional flow versus using arrows to indicate a relationship between two items or ideas (e.g. Figure 2.4.2). A System-level Depiction code was included for those responses, both written and drawn, that failed to describe either respiratory or circulatory components beyond a systems-level description (e.g. “circulatory system” alone was drawn or described, with no mention or depiction of finer-grained structures like the heart or capillaries). Additionally, drawn responses were coded as Slide Diagram if the drawing attempt was a reproduction (or attempt at reproduction) of a supplied drawing from that lecture’s slides or those of a previous lecture; see Appendix A for full coding structure.

2.4.4 Coherence

Coherence (COH) measurements were an assessment on a 0 – 3 scale of how seamless a student explanation is in response to the prompt, and were scored after Key Concepts, Alternative Conceptions, and Representation codes were assessed; however, as they were whole-notecard assessments like DoV and DD scores, they were included in a box on the upper right or left of the notecard, space depending. Low COH scores indicated gaps in student thinking and/or the presence of alternative conceptions or inaccuracies that resulted in the response profoundly deviating from expert-like sequencing of thoughts and ideas. Alternatively, students could have partial responses that contained correct ideas, but lacked linking phrases or diagrammatic elements between these ideas. Middle COH scores may have had minor gaps in thinking and/or alternative conceptions; extensive connections and relationships between scientifically acceptable ideas with logical sequencing resulted in high COH scores.

2.4.5 Expert Score

After coding for all of the above codes, notecards were again holistically examined to determine where on a novice-to-expert scale the student response fell and given an Expert Score (ES). Considerations for determining an ES included the Coherence score, whether students included one or more Key Concepts, and how well or accurately those Key Concepts were described and contextualized. Although Coherence scores were important in determining expertise, flawed (but “gapless”) reasoning could result in a higher Coherence score but still fall short of an expert-like response; the Expert Score thus also assessed qualities that were content- and discipline-specific, such as the interconnectedness of ideas, thinking across biological scales, and systems-based thinking. Expert Score codes are summarized in Table 2.4.1 below as well as in Appendix A.

Score	Expert Score (ES) description
0	Prompt not addressed; comparison may be made instead of linkage described
1 (Novice)	Only respiratory or circulatory system is described or depicted at some level, but not both Both components (circulatory and respiratory) are described or depicted in some way, but there is no linkage or clear understanding of linkage described that addresses the prompt specifically
2 (Medium)	Connection between respiratory and circulatory system shown to some degree, but may not be described or depicted <i>explicitly</i> with capillaries, physical proximity/integration, and/or that blood is the means of transporting oxygen over long distances that cannot be covered by diffusion alone. Countercurrent exchange may be depicted but not contextualized.
3 (Expert)	Linkage between systems clearly and thoroughly described/depicted AND contextualization of countercurrent exchange in fish examples (if lungs depicted, not necessary to include)

Table 2.4.1 Description of Expert Score (ES) categories. Expert scores were assigned after all other coding was complete.

2.4.6 Alternative Conceptions

For the purpose of this study, alternative conceptions (ACs) were considered responses that were either opposed to or inadequately explained the processes at hand as compared to a scientifically accurate representation or description (i.e. the “expert” response). The list of Alternative Conceptions was initially drafted before any notecards were examined, with best guesses as to what ideas students would potentially convey, but was edited and refined following notecard analysis to collapse some categories and add others. The final draft of ACs had some codes that specifically referred to inaccuracies, such as how the circulatory system was depicted or described (e.g. Figure 2.4.3), whereas others addressed a lack of clarity in responses (Unclear intent; Terminology) that was required in order to deem a sufficient understanding of the subject. Analysis of written responses from students also produced a third major type of AC, Comparison, in which the question prompt was not interpreted correctly and students compared similarities between the two systems rather than addressing their linkage. Student responses were

also assessed and coded for an alternative conception code if they included gratuitous diagrams or drawings that did not clarify or improve their response (Unessential Drawing), or if they missed a critical component of the system(s) necessary for a complete response (Key Component Absent). The list of AC codes and their descriptions can be found in Appendix A.

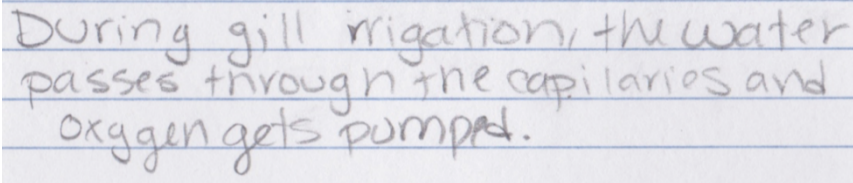
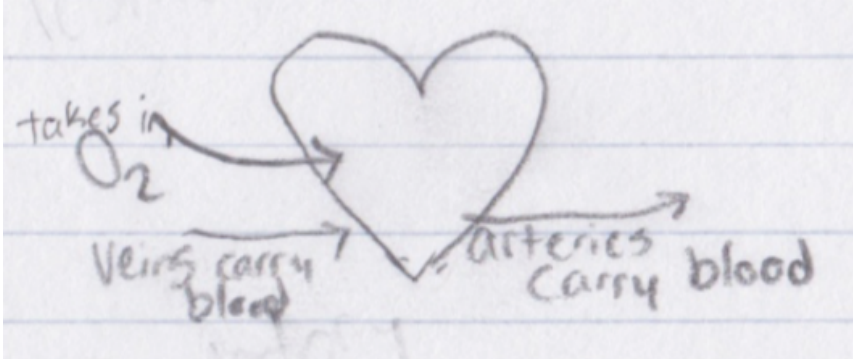
- a. 
- b. 

Figure 2.4.3 Student examples of the “Circulatory Inaccuracy” Alternative Conception (AC) code. Both student responses here earned the AC code of “Circulatory Inaccuracy.” The written response (a) incorrectly described *water* as the fluid passing through capillaries in the gills, rather than *blood* as the fluid within fish circulatory systems. The drawn response (b) attributes oxygen intake directly to the heart itself instead of naming or describing any component of a respiratory system (e.g. lungs or gills), which are the structures that are actually responsibly for oxygen intake.

a.

17R&C_457

DoV: 0 **COH: 3**

DD: N/A

KC: Circulatory and respiratory systems are physically integrated
Evidence: writing alone

KC: Mass transport
Evidence: writing alone

Representation (process): Gas exchange/diffusion

Representation (structure): Lungs/gills, written molecule, capillaries

ES: 2

b.

17R&C_409

DD: N/A **DoV: 0**

COH: 1

AC: Comparison

Representation (process): Gas exchange/diffusion

Representation (structure): Capillaries

ES: 0

AC: Unclear intent

Figure 2.4.4. Examples of coded 2017 student written responses. The top notecard (a) reads: “In the gills, oxygenated water is in close contact with the capillaries. From there, the O₂ in the water diffuses into the capillaries where the O₂ can be transported throughout the body via the circulatory system.” The bottom notecard (b) reads: “Respiratory and circulatory systems are closely related because they both deal with exchange/diffusion in some ways as well as capillaries.” Boxes for Coherence (COH), Degree of Visualization (DoV), Depth of Drawing (DD), and Expert Score (ES) codes are arranged as shown in all notecards when possible. Specific representations are boxed separately, as well as language/wording that captures Key Concept (KC) code ideas. Notecard (a) received an ES of 2, as countercurrent exchange is relevant to this example (gills) but was not mentioned or described. Notecard (b) received an ES of 0, as the response compared the two systems and thus did not adequately address the prompt.

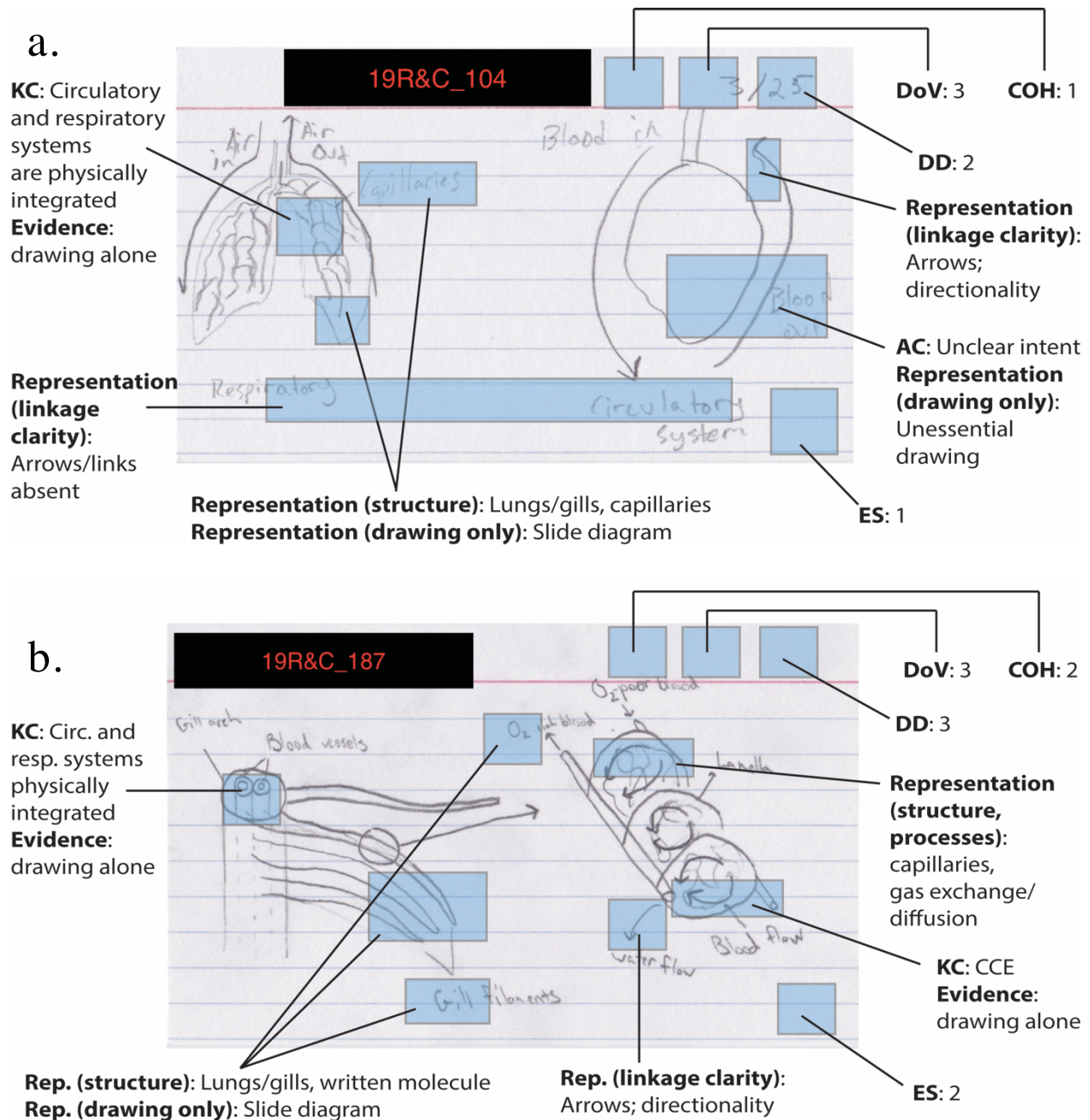


Figure 2.4.5. Examples of coded 2019 student drawn responses. The “Rep.” abbreviation is used in places for the “Representations” parent code. The top example (a) depicted both systems to a degree, but the diagram of the circulatory system on the right was both unclear as far as structures depicted, arrows showing blood flow, and did not support the idea of the linkage between the two systems. Although this student included capillaries in the lungs to demonstrate an understanding of a Key Concept (physical integration), the rest of the card was disjointed and thus the overall ES was 1. The bottom notecard (b) showed thinking across different scales to depict countercurrent exchange in the gill filaments, but failed to contextualize within an organism-scale framework to depict mass transport, and thus received an ES of 2.

2.5 Data Analysis

2.5.1 Sample pooling

Once all cards were coded according to the coding structure described above, data were downloaded from Dedoose for analysis in Excel. Data were downloaded both as code count per media (“Code Application” under “Media Charts” in the Analysis function of Dedoose) as well as code presence/absence per media (“Code Presence” under “Media Charts”); given that media were named according to the year of data collection, these data sets preserved the ability to discern which media came from which year. For analysis purposes, 2017 and 2018 were pooled into a singular “Written” category, as all but one of the 137 responses between these two years were written; 2019 responses were grouped into a “Drawn” category, as the prompt for that year explicitly asked students to draw their responses. Although this meant that there were more written than drawn responses, the larger sample size obtained by pooling the two written years’ data maximized the likelihood of seeing a less biased code frequencies that could result from smaller sample sizes (e.g. the $n = 58$ responses in 2017 alone). Additionally, lecture materials and delivery were nearly identical across all three years; therefore, it was assumed that 2017 and 2018 data were comparable.

2.5.2 Frequency calculations

Due to the discrepancy in sample sizes between written and drawn responses, code frequencies were presented as percentages of total response numbers to account for this difference. For most codes, the number of times the code appeared in a response category served as a proxy for the number of cards containing that code, as many of the codes were only ever applied once to a single notecard (e.g. Capillaries would only be coded once even if the word or diagram appeared multiple times on a single response). For others that could potentially appear

more than once on a card, such as the Alternative Conception code for Unclear Intent, frequencies were still calculated as the number of code occurrences out of total number of responses, but were noted when they occurred more than once per card. Likewise, for Key Concept (KC) codes, it was possible that multiple KCs showed up on a single card and thus a card may be “counted” twice when calculating frequency of code appearance between KC code categories. Certain analyses, such as comparisons of co-occurring codes like Key Concept-Expert Score (ES) codes, had frequencies that were calculated out of number of responses per category (e.g. the number of drawn ES 1 cards that also contained the Mass Transport KC code out of all drawn ES 1 responses). The method of frequency calculation was noted in all figure axes labels and captions, as appropriate. When card count data made more sense, such as for low-occurring code counts, this was also noted in figures and accompanying text.

2.5.3 Chi-square analysis

When comparing categorical written versus drawn code frequencies, two-way Chi-square analysis was performed to determine if differences in code frequencies between the two response categories were significant at $\alpha = 0.05$. Chi-square analyses were performed in Excel by comparing the frequencies of code occurrence (i.e. the percent of responses in which the code was present, versus the percent of responses in which the code was absent) in the written versus drawn categories.

CHAPTER 3

RESULTS

The purpose of this study was to assess the utility of including more drawing-to-learn activities in a marine science classroom by analyzing whether and how students communicated knowledge differently on a comparative anatomy topic when asked to draw (versus write) their response. If differences between the quantity of ideas (e.g. how many discrete concepts are presented) as well as the quality of response (e.g. how well students describe the integrated nature of the concepts in a scientifically accurate and coherent manner) exist between the two response types, then this gives valuable insight into how instructors should be posing formative assessment prompts on similar topics to better elicit student ideas and thus produce better data for instructor feedback. To that end, analyzing (1) the actual pieces of information or ideas that students communicated on their notecards (i.e. Key Concepts and Representations), (2) the expertise and coherence of the response, and (3) the nature of alternative conceptions present (if any) provides a means to compare and assess the two response types. The following results address each of these pieces (and thus, Research Questions 1 – 3) in turn.

3.1 Student communication of knowledge

Given the methodology employed in this study, addressing Research Question 1 (how students communicate their content knowledge in response to a given prompt) was achieved through the analysis of the coding structure constructed and applied to the notecard responses. Since follow-up interviews were not conducted to tease out meaning behind notecard responses, the coding scheme was utilized with the understanding that responses could only be interpreted as they appeared; that is, what students chose to put on the notecards simply demonstrated their knowledge on the topic of circulatory and respiratory integration in marine animals at that point

in time in the context of a given response category. No conjecture could therefore be made as to whether the notecard response truly captured the complete understanding of the topic on a student-by-student basis. Despite this, the notecard analysis by coding allowed for a detailed examination of *communication* of ideas – be they written or drawn – which still offers a valuable insight into how questions should be asked in order to best allow students to demonstrate the knowledge they *do* possess on a topic. In particular, the Key Concept and Representation code categories were especially helpful to identify and classify what kinds of knowledge, and the depth of that knowledge, that students were able to communicate in their notecard responses.

3.1.1 Key Concept codes

The number of Key Concepts used in student responses varied depending on response type, with drawn responses consistently displaying a higher frequency of responses that contained at least one Key Concept (KC). Neither the Written nor the Drawn category of responses contained any cards that had all three Key Concepts coded. Of the written response types, nearly half (44.9%) lacked any type of KC code, compared to only 25.6% of drawn responses (Figure 3.1.1; “0” category). The frequency of cards that contained at least one KC code between the two response types was much closer, at 47.8% for written responses and 56.4% for drawn. Finally, drawn responses showed a higher frequency of cards that demonstrated two KC codes (19.0%), versus only 6.6% of written responses coded for two KCs. A two-way Chi-square test ($\chi^2 = 12.758$, $df = 2$, $p < 0.01$) indicated that these observed frequencies of number of KCs were indeed significantly different between the Written and Drawn categories.

The frequencies with which Key Concepts appeared within a response category varied between written and drawn as well. For written responses, “mass transport” (MT) and “circulatory and respiratory are physically integrated” (CRI) appeared at nearly identical

frequencies of 25.5% of responses and 26.3%, respectively (Figure 3.1.2). “Countercurrent exchange” (CCE) only appeared in 9.5% of written responses. Like the Written category, CCE in drawn responses appeared the least frequently out of all three Key Concepts, representing only 12.7% of drawn notecards that had at least one KC coded. In the Drawn category, however, there was a larger difference observed between the frequencies with which MT and CRI appeared; MT was indicated in 45.5% of responses, and CRI in 34.5%. Although this may seem to suggest that mass transport is potentially more easily communicated in drawn form over written, it was important to examine drawn notecards in more detail to determine if this was in fact the case. To that end, the “Use of Evidence” code allowed for a closer inspection of drawn responses to obtain a more nuanced understanding of these Key Concept type frequencies.

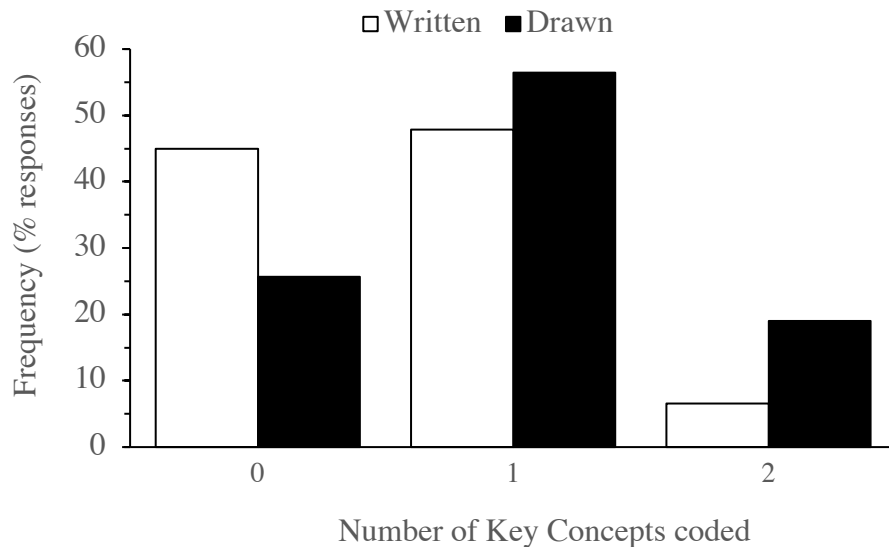


Figure 3.1.1. Number of Key Concepts coded by response type. Percent of cards per response type (written versus drawn) with 0, 1, or 2 Key Concepts coded (note: no card in either response type had all of the 3 possible Key Concepts coded). A two-way Chi-square test ($\chi^2 = 12.758$, $df = 2$, $p < 0.01$) indicated the number of Key Concepts coded depended upon response type.

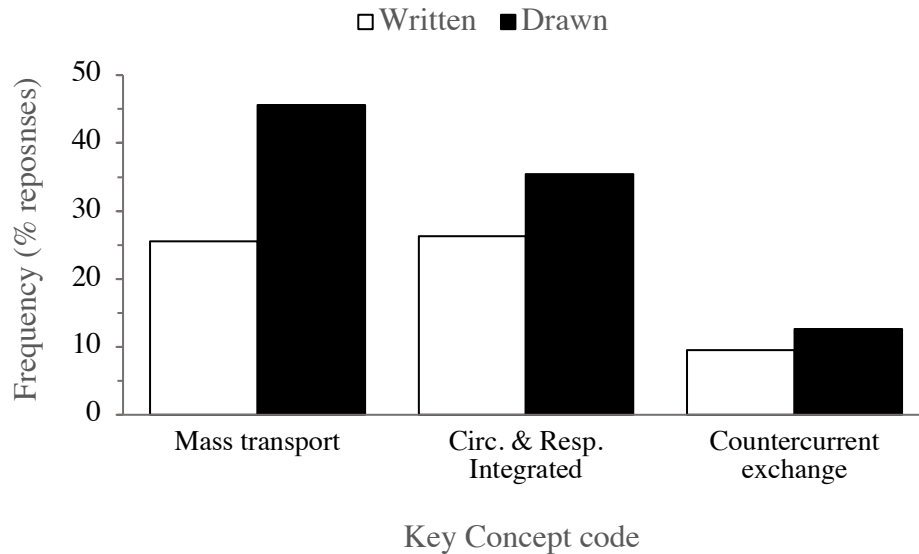


Figure 3.1.2. Type of Key Concept code by response type. Percentage of cards per response type (written versus drawn) that were coded for each Key Concept code. Data reflects total code counts and thus one card’s data may appear in two different categories if two Key Concepts were coded on a single card.

The Use of Evidence code was co-coded with any Key Concept code to determine how the Key Concepts were depicted on notecards, particularly to determine if students were actually using drawing and diagramming to convey crucial information in 2019 (“Drawn” response category); these results are summarized in Figure 3.1.3. In considering the data depicted in Figure 3.1.2, this was critical to determine if the difference observed between the frequencies of mass transport and circulatory – respiratory integration Key Concepts in drawn responses (as compared to the nearly identical frequencies observed in the Written category for these two KCs) was actually due to mass transport being depicted in drawn form more often. Drawn data was scrutinized in more detail to better evaluate the utility of this drawing-to-learn exercise in this regard.

For the drawn category of responses, three of the five Use of Evidence codes were analyzed: “Drawing alone,” “Both written and drawn,” and “Writing alone when asked to draw,” the latter of which was unique to coding 2019 notecard responses. “Writing alone” and

“Flowchart” Use of Evidence codes did not co-occur with any Key Concepts in 2019, and thus were excluded from analysis. Students, on the whole, appeared to heed the prompt instructions to draw or diagram their responses when communicating their knowledge about ideas critical to circulatory-respiratory linkage. For responses coded for mass transport, 58.3% contained some kind of diagrammatic or drawn support for this Key Concept; amongst these responses, over half relied entirely on drawing to communicate mass transport (33.3% of total for the category; see Figure 3.1.3). For circulatory – respiratory integration (CRI), the responses relied on drawing even more heavily: 60.7% of all cards coded for this KC used *only* drawing to communicate this idea (Figure 3.1.3). Strictly written responses were entirely absent from cards coded for the countercurrent exchange (CCE) Key Concept; 100% of responses relied on drawing in some form, even though the majority (70.0%) relied on a balance of both written and drawn evidence. Taking these results together with Figure 3.1.2, it appears then that the higher percentage of drawn responses that communicated mass transport (35 out of 79 cards, 45.5%) over the integration of circulatory and respiratory systems (28 out of 79 cards, 35.4%) cannot necessarily be attributed to students actually *drawing* mass transport on their cards in 2019. Indeed, 41.7% of cards in the Drawn category that conveyed mass transport actually relied strictly on written evidence (i.e. explanatory sentences or captions) to communicate this Key Concept.

Interestingly, despite the prompt in 2017 and 2018 not explicitly demanding a written answer from students, the overwhelming majority chose writing alone to communicate their knowledge on the Key Concepts. Students described MT, CRI, and CCE through writing alone in these two years in 97.1%, 100%, and 92.3% of all responses that contained each Key Concept code, respectively (data not shown). Only one response each accounted for the remaining percentage of responses that chose to incorporate drawing in some form for MT (2.9% of 35 total

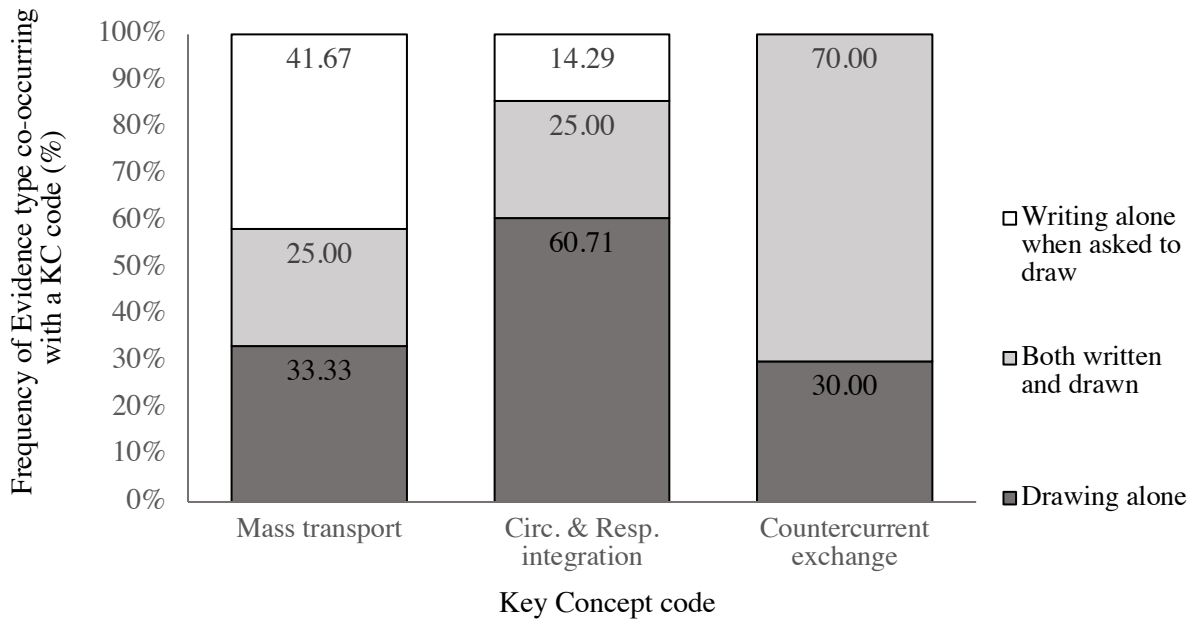


Figure 3.1.3. Use of Evidence codes per Key Concept for drawn responses. Frequencies were calculated as percentages of total cards coded for a particular Key Concept that used each type of evidence. Key Concepts were mass transport (MT; $n = 35$ cards), circulatory and respiratory systems are physically integrated (CRI; $n = 28$ cards), and countercurrent exchange (CCE; $n = 10$ cards).

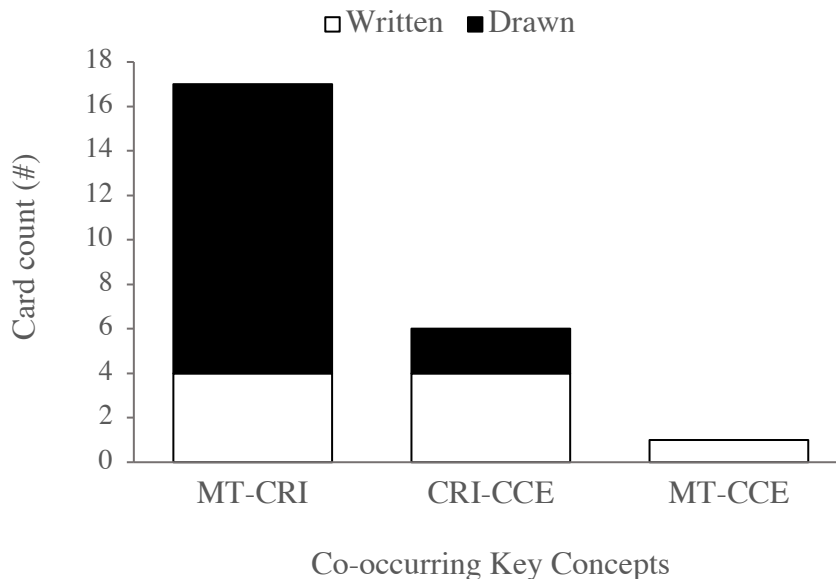


Figure 3.1.4. Card count data for co-occurring Key Concepts. Data comes from all cards that had two Key Concepts coded (“2” category from Figure 3.1.1). MT = mass transport, CRI = circulatory and respiratory systems are physically integrated, and CCE = countercurrent exchange. $N = 15$ cards for drawn responses, $n = 9$ cards for written.

responses) or CCE (7.7% of 13 total responses). Without explicit instruction and/or prompting to draw, students default to writing their responses, despite the fact that either drawing alone (as for CRI) or a combination of drawing and writing (as for CCE) may better communicate and support student thinking (Figure 3.1.3).

As described above, few cards in either response category captured two Key Concepts (Figure 3.1.1, “2” category): only 9 out of the 137 responses in the Written category (6.6%) and 15 out of the 79 Drawn (19.0%). Despite these small sample sizes, certain Key Concepts co-occurred more frequently than others. Beyond providing insight into what ideas students considered to be relevant in their response to the question prompt, co-occurrence of Key Concepts could also indicate that students are connecting concepts that were presented in different lectures and/or as part of different sub-topics within the overall lecture framework (e.g. a specific set of slides that discuss open versus closed circulatory systems in great depth one day, and another set that compare lungs versus gills) and synthesizing these discrete pieces of information into a coherent mental model.

Total card count data (rather than frequency of occurrence) showed that MT and CRI were much more likely to co-occur in drawn responses than either CRI – CCE or MT – CCE (Figure 3.1.4). Written responses, however, showed equally frequent card counts of co-occurrences of MT – CRI as CRI – CCE, with four responses appearing in each co-occurrence category. These results suggest that diagramming potentially facilitated students to make connections between the mass transport and circulatory – respiratory integration Key Concepts and thus communicate both ideas together more often. No drawn responses showed any co-occurrence of MT and CCE, and only a single written response fell under this category. Low instances of co-occurrence with CCE by either of the other Key Concept codes, however, was more likely due to overall low

occurrence of the CCE Key Concept code in all student responses, written or drawn, as seen in Figure 3.1.2.

3.1.2 Representation codes

Students in the written-response years generally used more holistic terminology when describing the linkage between the circulatory and respiratory systems, whereas students in the drawn-response year tended to describe components of each system in more detail. Certain key structures and processes that illustrate the linkage between the two systems, such as gas exchange, written molecules, and capillaries, were used to demonstrate student understanding equally between written and drawn responses (Figure 3.1.5; “ns” columns). Interestingly, apart from the “Capillaries” structure code that appeared roughly equally both response categories, the codes that occurred more frequently in drawn responses showed finer-grain details of either or both systems by naming particular sub-structures or networks of structures, rather than a more generalized system-level description or depiction (e.g. merely stating “the respiratory system,” rather than describing the lungs or gills). In particular, 58.2% of drawn responses depicted the heart in some capacity as compared to only 7.3% of written responses (Figure 3.1.5; $\chi^2 = 73.08$, $df = 1$, $p < 0.001$). Likewise, the percentage of drawn responses that depicted the lungs and/or gills (84.8%) was nearly double that of the written responses depicting these structures (43.8%) (Figure 3.1.5; $\chi^2 = 38.26$, $df = 1$, $p < 0.001$). Finally, the System-Level Depiction code appeared at a significantly higher frequency in written responses (33.6%) than in drawn (3.8%) (Figure 3.1.5; $\chi^2 = 25.33$, $df = 1$, $p < 0.001$).

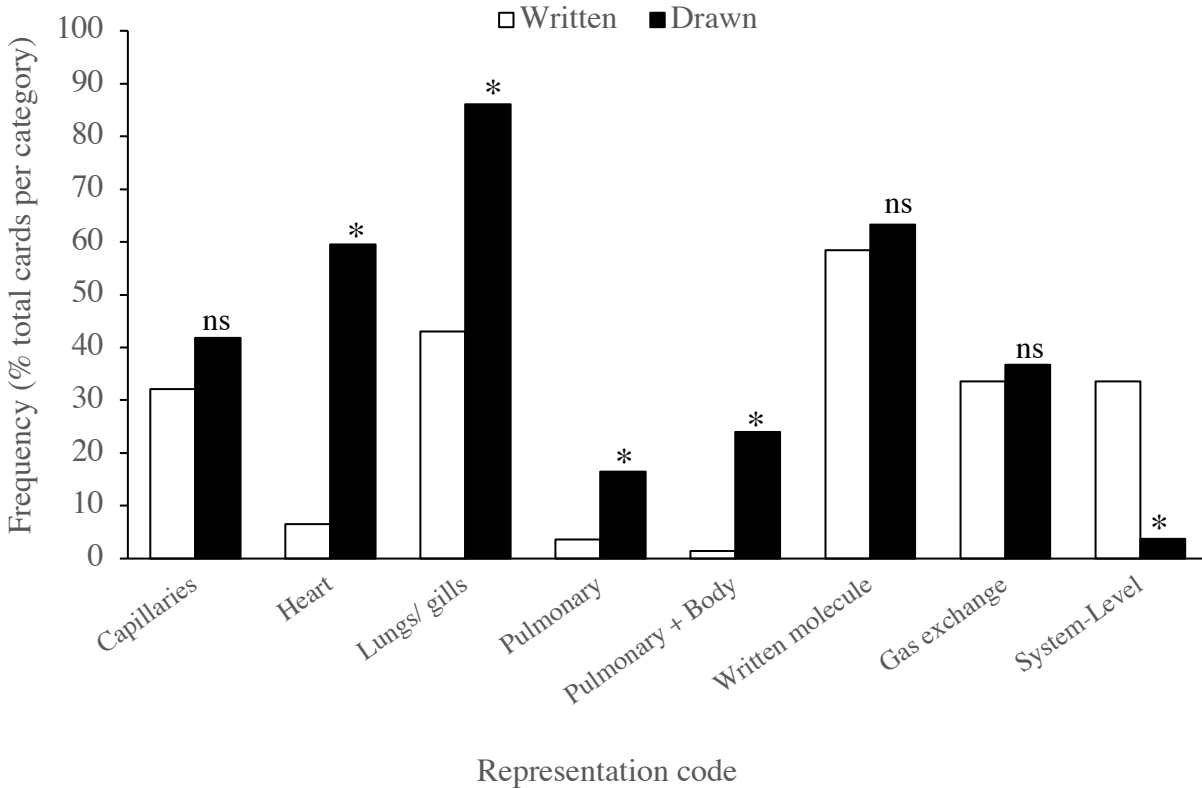


Figure 3.1.5. Representation codes frequencies by response type. Frequency of Representation code usage in Written (white) versus Drawn (black) responses. Stars (*) indicate a significant difference in frequency between written and drawn responses for that particular structure code (χ^2 , $df = 1$, $p < 0.01$), whereas “ns” indicates “no significance.”

3.2 Assessing novice versus expert responses

Addressing Research Question 2 (comparing written versus drawn responses across novice-to-expert scale) relied on assessing demonstrated knowledge (from the Key Concepts and Representations code data described above) as well as fluency of response (e.g. idea linkers that provided “gapless” thinking) in order to categorize responses as “novice,” “expert,” or somewhere in between. Additionally, responses needed to be assessed holistically by considering other important factors such as Alternative Conceptions codes, and whether these detracted enough from demonstrated knowledge to impact a student’s demonstration of expert-like thinking. After all these intersecting and overlapping code structures were considered, Expert

Scores provided a means by which all response types could be assessed along a novice-to-expert scale. Expert Score codes were therefore key to determine how these trends differed not only across response type, but also how these score categories co-occurred with other codes to provide a more complete characterization of each category.

3.2.1 Expert Score breakdown by response type

Expert Scores (ES) between the two response categories showed a tendency for drawn responses to score higher (Figure 3.2.1). No difference was revealed in written versus drawn response types for cards scoring an ES of 1 (“Novice” category); 38.7% and 38.0% of written and drawn responses, respectively, earned this score. Only two responses in the entire pool (one written, one drawn) received an ES of 3 (“Expert” category). However, drawn responses tended to be more likely to score a 2 (“Medium” category ES) as nearly half of all responses in this category (49.4%) received an ES of 2, versus only 36.5% of written responses. Accordingly, the frequency of written responses earning an ES of 0 (“Prompt not addressed” category) was greater than that of the drawn responses (24.1% versus 11.4%, respectively).

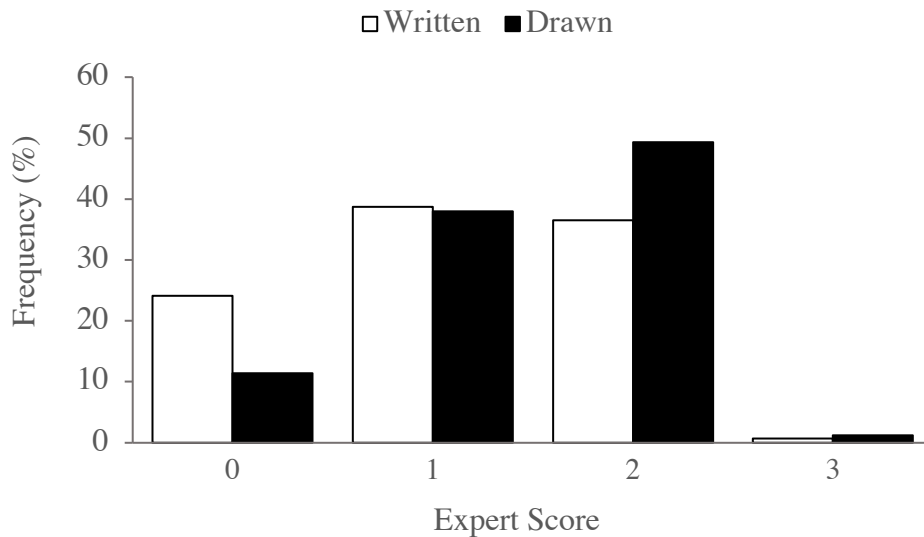


Figure 3.2.1. Frequency of Expert Scores (ES) between the two response categories. Written (white) versus Drawn (black). Percent values were calculated as the number of cards that received the particular ES out of the total number of cards per response category. ES 3 for both Written and Drawn only had one card each.

A closer examination of drawn responses with an ES of 0 ($n = 9$) revealed that only one did not have any drawing attempt whatsoever. Four of the cards had rough drawings indicating some kind of cycling of materials; interestingly, three of these four made some kind of comparison between the two systems (noted by the Comparison Alternative Conception code), as indicated by the example card in Figure 3.2.2. Similarly, in most of the written responses that received an ES of 0 ($n = 33$), the Comparison code appeared frequently (discussed in more detail in Section 3.4). Interestingly, however, there seemed to be verbal descriptions amongst these ES 0 cards that mirrored the cycling diagrams depicted in some of the ES 0 drawn responses. These written descriptions discussed moving materials throughout the body, which sometimes earned them a Key Concept code for mass transport, but often fell short on supporting details that were accurate or complete enough to earn a higher ES score. Descriptions included verbiage that directly or indirectly alluded to cyclical events, such as “respiratory systems go through a cycle,”

“they both uptake and release,” “both flow throughout the body,” and “they transport something in order for the organism to live.” As such, it appears that many students who missed the mark entirely in addressing the prompt either misinterpreted or misapplied mass transport of gasses throughout the body, oftentimes leaving their description of such transport at a very surface-level interpretation, or communicating that *both* systems moved gasses throughout the entire body. The remaining ES 0 cards, both written and drawn, that did not fall under these comparison/cycling themes generally had vague and/or inaccurate depictions of how the two systems were linked, with minimal (or inaccurate) diagramming or description to support the main ideas.

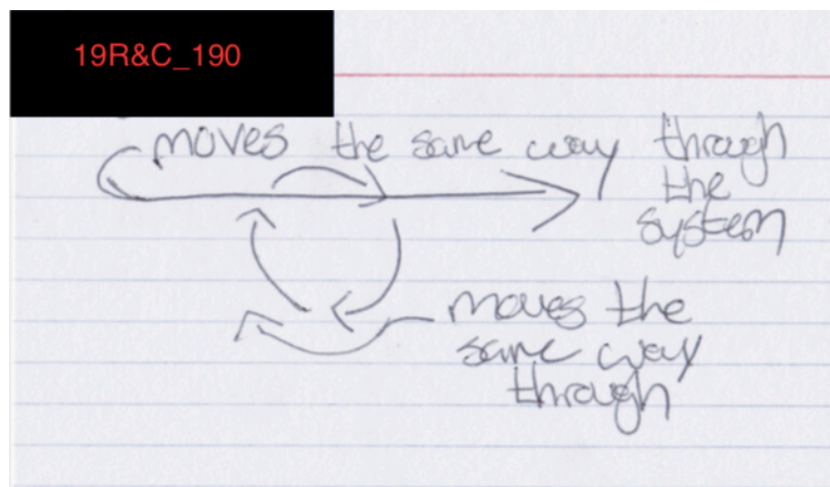


Figure 3.2.2. Sample of Expert Score (ES) of 0 in a drawn response. Sample of drawn response card eliciting an ES 0 that indicated cycling of materials in some way but also compared the two systems.

3.2.2 Key Concept co-occurrence with Expert Score

A closer examination of codes that co-occurred with each of the Expert Score code categories revealed deeper insights into the overall assessment of the utility of written versus drawn responses. Key Concept (KC) codes were considered to be critical to capturing an expert-like understanding of the content; as such, assessing how many and which KCs appeared in

responses for each ES score was important to understand how students across the novice-expert scale were utilizing key information in their answers, and whether this varied by response type.

Despite this importance of Key Concepts in conveying expert-like understanding of the integration between circulatory and respiratory systems, cards without any coded did not necessarily elicit an Expert Score of 0 (“prompt not addressed”). Of all cards that received an ES of 1, 58.5% of written responses and 40.0% of drawn responses had no Key Concepts coded (Figure 3.2.3, dark gray bars in ES 1 category); there were sufficient details that were communicated on these cards that may not have corresponded to particular Key Concepts, but demonstrated enough knowledge on the topic that warranted a novice score. However, most

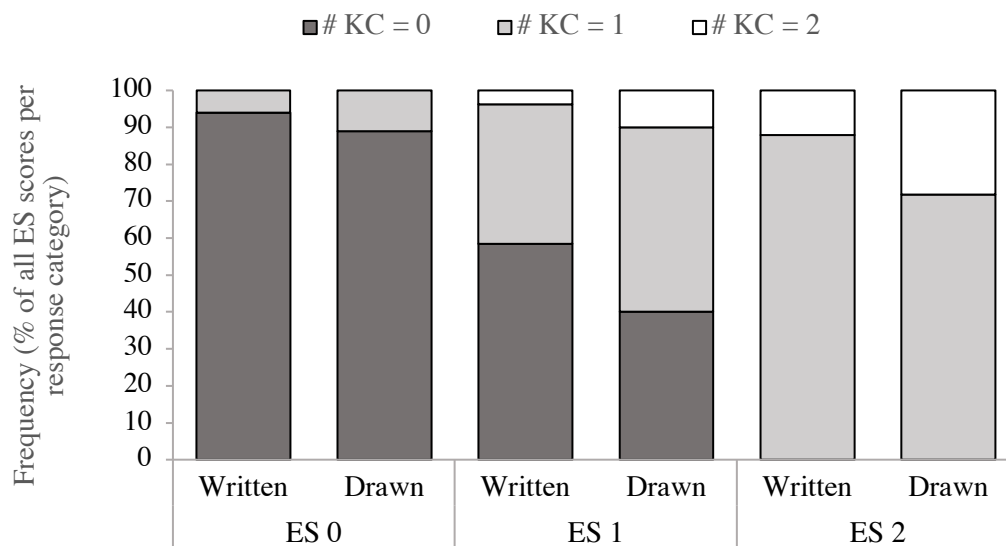


Figure 3.2.3. Frequencies of number of Key Concepts (KC) across Expert Scores (ES). For each response category, frequencies were calculated as the number of responses that had a given number of KCs (e.g. how many cards were coded for 1 Key Concept) out of all cards that received a given Expert Score. Sample sizes for each ES category are as follows, with written and drawn listed respectively: ES 0: $n = 33, 9$; ES 1: $n = 53, 30$; ES 2: $n = 50, 39$. The two cards (one written, one drawn) that received an ES of 3 are excluded from this figure.

cards that received at least one KC code fell under the ES categories of 1 (“novice”) or 2 (“medium”); only two written and one drawn response had one KC coded but still only scored an

ES of 0, representing 6.1% and 11.1% of all ES 0 cards, respectively (Figure 3.2.3, light gray bars in ES 0 category). In general, although capturing Key Concepts in a student response was not a prerequisite for earning a higher Expert Score, both written and drawn responses showed a trend in which greater numbers of KC codes translated to a correspondingly higher ES.

Next, responses across both categories that only had one KC code were examined ($n = 66$ written responses and $n = 44$ drawn) for the breakdown of what *type* of KC appeared. Neither written nor drawn responses had any cards with a single KC score an ES of 3 (“expert”); however, only describing a single KC did not preclude responses from earning an ES

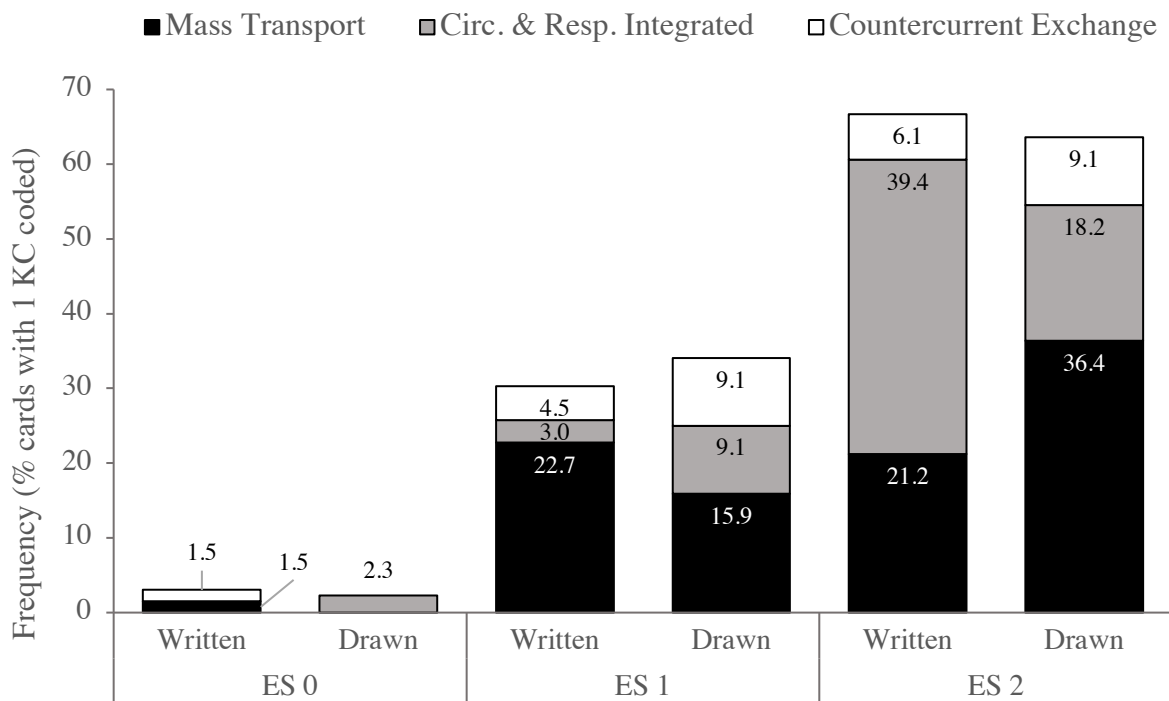


Figure 3.2.4. Frequency of each Key Concept (KC) code type by response category and Expert Score (ES). Values on the bar represent the percentages as calculated by the number of times a particular KC code appeared in the given response category + ES over the total number of responses per category that had 1 KC; $n = 66$ written and $n = 44$ drawn responses, respectively, that had only 1 KC coded.

of 2 (“medium”). Of these singular Key Concept cards that earned an ES 2, for both response types, mass transport (MT) and the integration of circulatory and respiratory systems (CRI) were

the two most common KC codes to appear, with countercurrent exchange (CCE) appearing in only 6.1% of the ES 2 written responses and 9.1% of the ES 2 drawn responses (Figure 3.2.4, ES 2 category). For written responses, the CRI code appeared to be a good predictor of higher ES scores, as only 3.0% of responses in the ES 1 category were coded for CRI, versus 39.4% of the ES 2 cards. Drawn responses had a rough doubling of frequencies for both the MT and CRI Key Concept codes between the 1 and 2 Expert Score categories, with the frequency of the MT code increasing from 15.9% to 36.4% from ES 1 to ES 2 and the CRI code increasing from 9.1% of responses to 18.2%. As such, it did not appear that either the MT or CRI code were good predictors of Expert Score for Drawn responses with a single Key Concept.

3.2.3 Expert Score co-occurrence with drawing-specific codes

Certain codes, such as Depth of Drawing (DD), Degree of Visualization (DoV), and whether the drawing produced was an attempt to reproduce lecture materials (“Slide diagram” under Representations parent code) were unique to drawn responses from 2019. This data was analyzed in the context of Expert Score (ES) data to better understand how, if at all, ES may have been informed by aspects of drawing in particular.

Depth of Drawing (DD) helped elucidate what type of drawing students created in response to the prompt, be it a baseline drawing that showed generic circuit-like loops between circulatory and respiratory organs and structures (DD 1), or more in-depth depictions such as the close investment of capillaries in gills and lungs (DD 2) or countercurrent exchange (DD 3). Accordingly, as ES increased, a greater representation of higher DD scores also increased (Figure 3.2.5); the percentage of responses that earned a DD score of 2 increased from 1.3% in the ES 1 category to 15.2% in the ES 2 category. Likewise, the frequency of DD 3 scores increased slightly from 2.5% to 7.6% across ES 1 to ES 2. Although DD 3 represented a higher-

order process (countercurrent exchange), because of the need to contextualize this process both on the organ (gills) and organism (fish) levels to demonstrate true expert-like knowledge and/or understanding, cards that received a DD score of 3 did not necessarily earn higher Expert Scores.

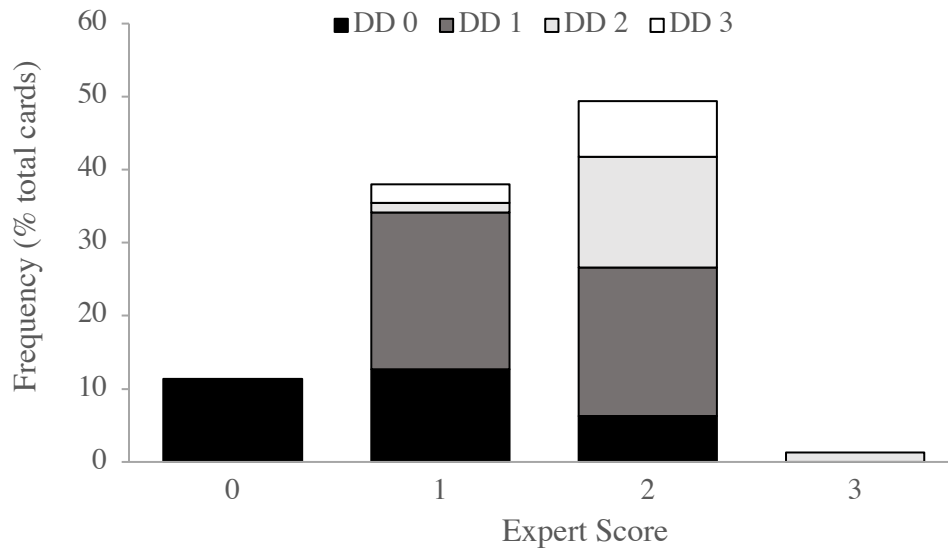


Figure 3.2.5. Depth of Drawing (DD) scores per Expert Score category. Frequency calculated out of the total number of drawn responses ($n = 79$ cards).

Degree of Visualization (DoV) scores were assessed to determine how much diagramming and drawing students actually incorporated into their responses in 2019 (data not shown). Few students, when prompted to draw, failed to draw anything; only 7 out of the 79 responses (8.7%) received DoV scores of 0, which represented strictly text/verbal responses with no accompanying visuals. Despite a lack of any visualizations, these seven responses spanned three categories of Expert Scores, with one response scored as an ES 0 and three each scored as ES 1 and ES 2, respectively; therefore, higher DoV scores did not appear to correlate with higher Expert Scores. Cards with a Degree of Visualization score from 1 – 3 were split evenly across the “novice” (1) and “medium” (2) Expert Score categories: 26.6% of all cards within each DoV score category (1, 2, or 3) received an ES 1 or ES 2 ranking. The single response from 2019 that

received an ES of 3 only had a DoV score of 2. Taken together, no trends were observed in comparing DoV scores across ES categories; it appears instead that the *quality* of diagramming (indicated by the Depth of Drawing scores, discussed above) outweighed the *quantity* of diagramming or drawing when determining expertise of knowledge for 2019 responses.

Expert Scores were also cross-compared with whether or not the response was coded for “Slide Diagram,” which indicated that the student drawn response was an attempt to capture or recreate a diagram or image that had been displayed in the lecture slides. Out of all 79 drawn responses, 23 cards (29.1%) were coded for “Slide Diagram,” all of which fell under either an ES of 1 or 2. Seven of these responses were scored at an ES 1 (representing 23.3% of all drawn ES 1 cards), and the remaining sixteen responses were scored at an ES 2 (representing 41.0% of all drawn ES 2 cards). As such, it appears that a recreation of a slide diagram from the lecture, even one that was potentially highly detailed and/or conveyed a higher-order process such as countercurrent exchange, did not necessarily translate to higher Expert Scores in drawn responses.

3.3 Alternative conceptions in novice responses

The third research question (RQ3) sought to determine whether the nature of alternative conceptions differed between written versus drawn responses; specific child codes under the parent Alternative Conception (AC) code category helped uncover any differences. Additionally, Alternative Conception codes were cross-referenced with Expert Scores to determine if written and drawn responses showed similar trends in the type of ACs that appeared as responses earned higher Expert Scores. In comparing each AC code in written versus drawn responses, a few notable differences were observed. Three AC codes in particular – “Vitalistic” Response, Comparison, and Terminology – were present at significantly higher frequencies in written

responses than in drawn (Figure 3.3.1, starred); $\chi^2 = 11.52, 17.89,$ and 19.47 for these three ACs, respectively ($df = 1, p < 0.001$). Two AC codes, Capillaries Absent and Unclear Intent, occurred at significantly higher frequencies in drawn responses over written (Figure 3.3.1); $\chi^2 = 4.63$ and 4.41 for Capillaries Absent and Unclear Intent, respectively ($df = 1, p < 0.05$). The remaining AC codes did not differ significantly between written and drawn responses.

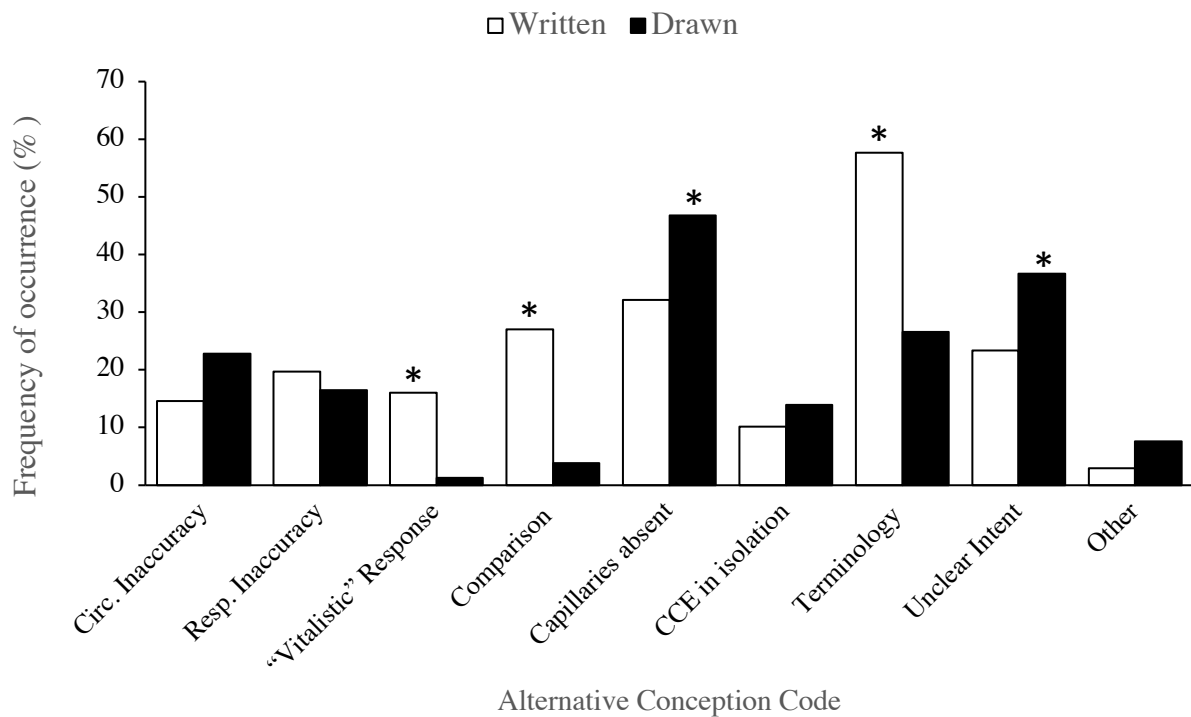


Figure 3.3.1. Frequency of Alternative Conception (AC) codes. Frequencies were calculated out of total responses for both written (white) and drawn (black) response categories. Stars indicate significant differences between written and drawn frequencies for that particular AC; “Vitalistic” Response ($\chi^2 = 11.52$), Comparison ($\chi^2 = 17.89$), and Terminology ($\chi^2 = 19.47$) AC codes occurred at significantly higher frequencies in written than drawn responses ($df = 1, p < 0.001$ for these three codes). Capillaries Absent ($\chi^2 = 4.63$) and Unclear Intent ($\chi^2 = 4.41$) codes occurred more frequently in drawn responses than written ($df = 1, p < 0.05$ for these codes).

To determine how Alternative Conception codes varied across novice responses, responses that were coded for Expert Scores of 0 or 1 (“prompt not addressed,” and “novice”) were further analyzed by which ACs were present in across written and drawn categories. ES 2

data (“medium”) was included to be able to determine if trends observed in AC code data were unique to novice responses, or if more expert-like responses also saw similar frequencies in ACs. Total AC code counts were tabulated for each response type and Expert Score (Table 3.3.1) in order to calculate frequency data for each AC out of total ACs coded.

Drawn Response AC Code Counts	Circ. Inaccuracy	“ Vitalistic” Response	Comparison	Capillaries absent	CCE depicted in isolation	Respiratory Inaccuracy	Terminology	Unclear Intent	Other	TOTALS
ES 0	2	0	2	0	2	3	6	6	1	22
ES 1	10	1	1	19	5	6	12	15	4	73
ES 2	7	0	0	18	4	4	6	12	1	52
Written Response AC Code Counts										
ES 0	4	11	18	1	4	10	28	20	3	99
ES 1	13	11	19	28	6	12	46	16	1	152
ES 2	3	0	1	15	4	5	23	1	0	52

Table 3.3.1. Raw count data for Alternative Conception (AC) codes by Expert Scores (ES) per response category. Total AC code counts are tallied at the end of each row, which were used to calculate frequencies of individual AC codes out of all ACs coded for each response type and ES category.

The bulk of AC codes fell under two categories, Unclear Intent (dark gray) and Terminology (diagonal thin stripes), which together accounted for 40-50% of AC codes for all novice responses, be they written, drawn, ES 0, or ES 1 (Figure 3.3.2). However, the total percentage of AC codes that these two ACs accounted for dropped steadily as Expert Scores increased, and accounted for less of the total percentage in drawn responses versus written in ES 1 and ES 2 categories. Interestingly, the Capillaries Absent AC code was virtually nonexistent in the ES 0 written and drawn responses, but appeared at a much higher frequencies in the ES 1 category (1.0% versus 18.4% for written; 0% versus 26.0% for drawn). Similarly, both written

and drawn responses that received ES 2 scores had higher code frequencies for Capillaries Absent, at 28.8% and 34.6% for written and drawn, respectively. The remaining AC code frequencies did not appear to differ to nearly as great a degree between Expert Scores or Written/Drawn categories, with the exception of the “Vitalistic” and Comparison AC codes, as discussed previously.

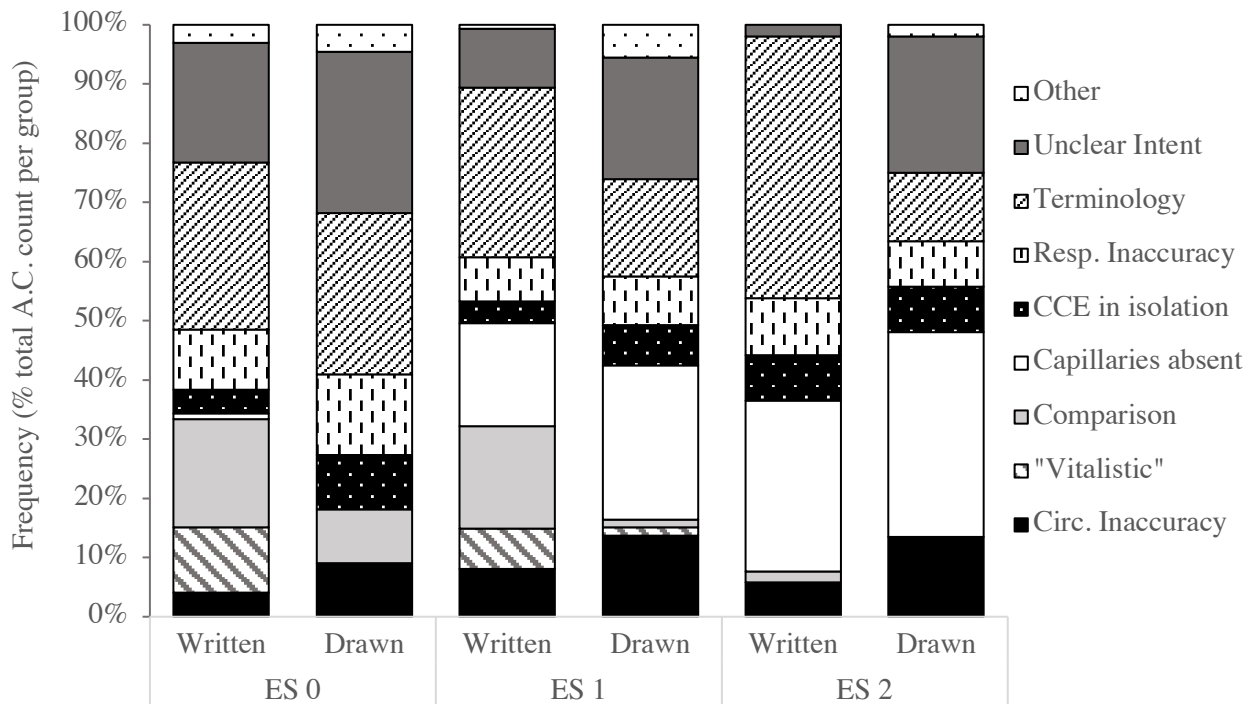


Figure 3.3.2. Alternative Conception code frequency by response category amongst novice responses. Novice score categories – Expert Score (ES) 0 and 1 – were compared to the medium category (ES 2). Frequencies were calculated as percentage of total Alternative Conception (AC) code counts for each response type and ES category. ACs appear in the bars, from top to bottom, in the same order as listed in the legend. “Respiratory” and “Circulatory” have been abbreviated as “Resp.” and “Circ.” for readability. Exceptions include the following columns for which certain ACs never appeared (listed as Data Column, ACs): ES 0/Drawn, Capillaries Absent and “Vitalistic”; ES 2/Written, Other and “Vitalistic”; ES 2/Drawn, Comparison and “Vitalistic.”

CHAPTER 4

DISCUSSION

The research questions posed in this study sought to articulate, compare, and contrast the differences between written and drawn responses to a formative assessment prompt in an undergraduate marine science course. With the goals of *Vision and Change* (AAAS, 2011) in mind, this analysis was an exploratory study into the efficacy and utility of implementing one type of drawing-to-learn strategy (diagramming spatial-visual comparative anatomy subject matter) to better understand from an instructional standpoint what information students deemed relevant and important in communicating their responses to a question prompt. The specific concepts or ideas that students presented in their responses – as determined through Key Concept and Representation coding data (RQ1) – gave insight into what students determined was relevant and important to address the prompt, particularly when viewed through the Selection – Organization – Integration (SOI) framework (Mayer, 2014). By comparing written and drawn coding data to address this research question, particular Key Concepts (if any were communicated) and/or Representations could be assumed to be more salient to students in the SOI process; thus, this circulatory – respiratory systems linkage question (and other prompts addressing similar visual-spatial subtopics) seemed to benefit from certain elements being drawn. Coherence and Expert Score data were used to determine the sophistication and accuracy of student responses between response categories (RQ2); here, too, the larger goals were to determine if prompting drawing within a constructed-response question captured different student ideas than previous written iterations. Ultimately, this uncovered a variety of useful instructional feedback, from greater depth of detail in student knowledge seen in drawn responses, to alternative conceptions in both response categories that must be explicitly

addressed through instruction. Along these lines, the final research question (RQ3) sought to specifically parse out the *types* of alternative conceptions that students communicate in written versus drawn form, in order to better understand how to facilitate their communication and (ultimately) understanding of key ideas. These results suggested that the more expert-like communication of ideas that was witnessed in drawn responses corresponded to similarly higher-order alternative conceptions, potentially due to the response medium (text/words versus pictures/diagrams) itself. The reasoning and implications for these findings are discussed below, with recommendations for future studies, as well as classroom practice.

4.1 Student knowledge: drawing may facilitate communication of information

A variety of interesting trends emerged from the Key Concept and Representation code data that demand closer inspection of the way in which instructors ask students questions, and the corresponding knowledge that students communicate in their responses. Although the data do not suggest that one response type over the other is universally better-suited for constructed-response question prompts such as the one in this study, it does appear that drawn and written responses each communicate different knowledge in different ways. Considering which type of response category appears to deliver more comprehensive student responses may be important for instructors when crafting assessment questions in order to best capture student knowledge of a given topic. Indeed, instructors could use formative assessments similar to the notecard questions in this study as opportunities to experiment with explicitly prompting written or drawn responses, thus giving critical insight into how summative assessment questions could be posed so as to give students the best opportunities to demonstrate their understanding of a topic on a case-by-case basis.

4.1.1 Visualization appears to facilitate cognitive coding and expression of certain Key Concept information

Key Concept (KC) data between response type indicated that drawn responses tended to include more Key Concepts than written (Figure 3.1.1). Accordingly, between the two response types, drawn responses had higher frequencies of each type of KC (Figure 3.1.2). A possible explanation for these trends could be due the more cognitively-demanding task of translating a visual diagram presented in slides to a verbal description. Several modes of representation are available for depicting scientific concepts, including the verbal mode (e.g. descriptions) and visual mode (e.g. graphs or diagrams); each of these modes requires practice and study in order to become familiar and fluent in their representations, thus the cognitive load for students, particularly novices, to switch between modes is demanding (Gilbert, 2005). Interestingly, it appeared that students tended to communicate certain Key Concepts in one mode over another, particularly when examining the drawn data in more detail. Mass transport, for example, was communicated in writing alone in 41.7% of the 2019 responses (Figure 3.1.3), despite the prompt explicitly asking students to draw their responses. Communicating this KC in writing involved students enacting the verbal mode, with students describing in some way that oxygen taken in by the respiratory system is moved around the body by the circulatory system. This often required the use of specific – yet fairly straightforward – language such as “transport,” or “throughout the body.”

The other two Key Concepts, however, appeared to be better communicated by drawing, given that responses communicating circulatory – respiratory integration (CRI) overwhelmingly used drawing alone (60.7%), and a further 25.0% used some combination of words *and* drawings; *all* countercurrent exchange (CCE) responses used drawing to some extent (Figure

3.1.3). Drawn responses, enacting the visual mode, often captured CRI and CCE by reproducing slide diagrams that did not require translating visual information into verbal information, and perhaps were easier for students to remember and/or were more salient to them when engaging in the SOI process while learning (Figure 4.1.1). Indeed, the cognitive effort involved for students

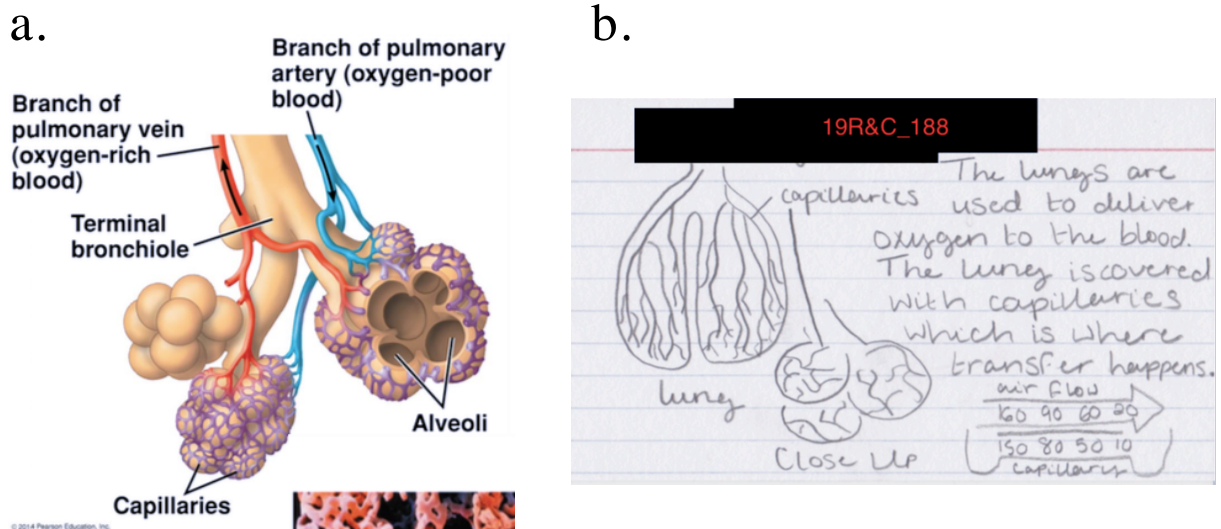


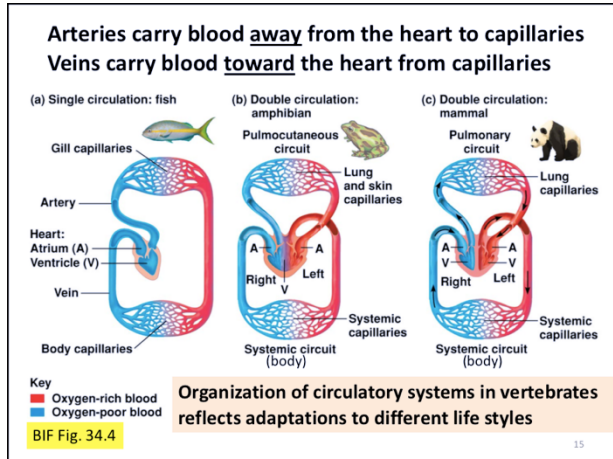
Figure 4.1.1 Circulatory – respiratory integration (CRI) Key Concept communicated in drawn form. A diagram (from Urry, Cain, Wasserman, Minorsky, & Reece, 2015) presented in the lecture slides (a) of capillaries being closely invested in the lung substructures (alveoli) is recaptured in the student diagram on the right (b) to communicate CRI.

to recall specific phrases to convey the CRI and CCE Key Concepts may be greater than being able to recall and reproduce a visual diagram, as in the first case students must “mode switch,” whereas in the second scenario students are demonstrating their knowledge in the same mode in which the information was initially presented. Furthermore, visual perception – i.e. the brain intaking and processing optical information – and the creation of a mental picture or imagery to codify information (akin to visually paraphrasing) are similar mental processes (Gilbert, 2005). The demonstration of more overall Key Concepts within drawn responses could therefore be the

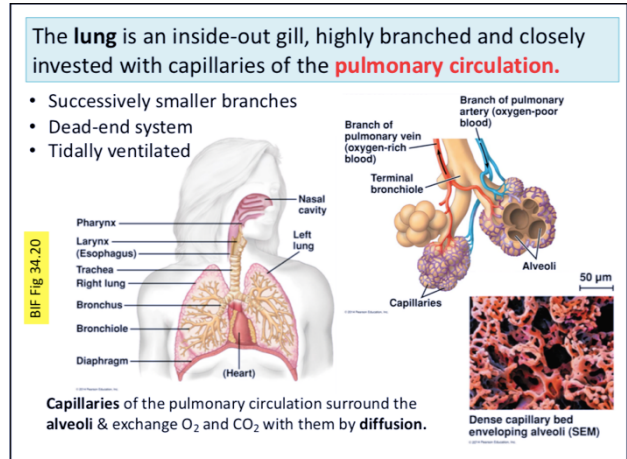
result of the particular knowledge being more easily codified visually, and thus easier to recall and reproduce in diagram form.

An additional important point to consider goes beyond considering whether concepts are presented visually or textually, but whether a diagram or written text is the *best* representation of the concept at hand. For the concepts that were identified as Key Concepts, perhaps, in fact, CRI and CCE are best conveyed and understood when they are displayed in a pictorial format. The adage “a picture is worth a thousand words” may be especially apt in this particular content area, and thus could account for a higher frequency of KC occurrence in 2019 responses as compared to the other two years. In the particular case of comparing marine respiratory and circulatory systems, diagrams were heavily utilized throughout the lectures leading up to the notecard question prompt in order to describe both large-scale and fine-grained structures and processes; see Figure 4.1.2 for examples. Although each of the slides in Figure 4.1.2 include text descriptions of the processes shown by the visual representations, the colorful diagrams were potentially (1) both easier to recall and reproduce, as suggested above, and/or (2) more accurate or thorough depictions of the Key Concepts chosen. The integration of circulatory and respiratory system and countercurrent exchange Key Concepts for this question prompt may have generally been easier for students to depict visually than verbally, particularly since “mode-switching” is cognitively demanding. Therefore, an important instructional implication may be that instructors should consider if what they consider to be “key knowledge” for students to demonstrate in their responses is visual or verbal in nature, and adjust their assessments accordingly to prompt the more natural mode of representation in student responses (e.g. explicit instructions to draw a response, or an option to draw or write).

a.



b.



c.

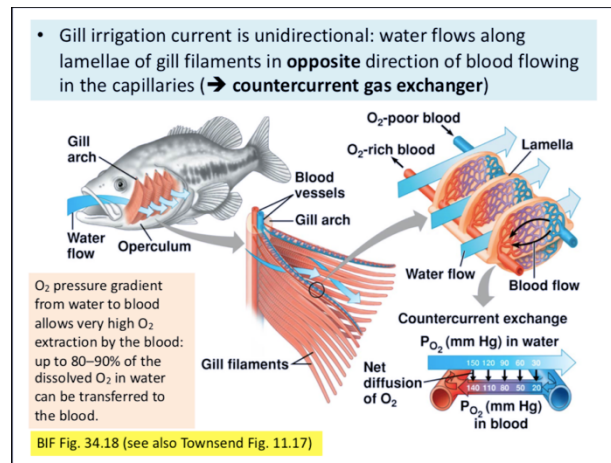


Figure 4.1.2 Visual examples of Key Concepts from lecture slides. Lecture slides were created by S. Lindsay and were essentially identical between years. Each slide example demonstrates the highly visual representations for each Key Concept, from mass transport (a), to the close integration of circulatory and respiratory systems (b), to countercurrent exchange (c). Images adapted from Townsend (2012) and Urry et al. (2015).

4.1.2 Depicting countercurrent exchange accurately is more difficult than other Key Concepts

Although drawn responses demonstrated higher numbers of Key Concepts, countercurrent exchange (CCE) appeared to be particularly demanding for students to communicate, be it through written or diagrammatic depictions. Indeed, the low co-occurrence of CCE with either of the other Key Concept codes was more likely due to the fact that countercurrent exchange was not only considered to be a “bonus” concept for students to include to achieve higher Expert and Coherence scores, but was also a more difficult concept to describe and/or depict than either CRI or MT. Although visualization of this Key Concept appeared to support student communication of this idea (Figure 3.1.3), overall students tended to struggle to accurately communicate CCE even when attempting to use diagrams, even though lecture slides contained visual representations of all three Key Concepts. For example, lecture slides contained multiple depictions of circulatory system “circuits” at varying levels of detail that could be referenced for depicting mass transport; see Figure 4.1.2a and Figure 4.1.3. Along these same lines, student responses were often able to depict the CRI (circulatory and respiratory systems are physically integrated) Key Concept with little effort. Verbally, it was typically sufficient for students to describe “capillaries in the lungs/gills” to earn this KC code. Visually, spiderweb-like capillaries overlaid on the lungs were easy enough for many students to depict on their circuit diagrams, and thus earn this KC code as well (e.g. Figure 4.1.1b).

In contrast, a quick examination of Figure 4.1.2c and Figure 4.1.4 can confirm the degree of detail required to adequately depict countercurrent exchange. Although Figure 4.1.2, Figure 4.1.3, and Figure 4.1.4 are diagrammatic examples of the Key Concepts, the language required to describe countercurrent exchange (e.g. “rete mirabile”) is also more challenging than that of

mass transport or circulatory-respiratory integration. Furthermore, countercurrent exchange can be used for more efficient gas exchange, as depicted in Figure 4.1.2c, or it can be used for heat exchange, as depicted in the examples in Figure 4.1.4, which can further complicate students' understanding of the process as it applies to the integration of circulatory and respiratory systems. Indeed, the student who submitted the notecard in Figure 4.1.1b included a recreation of a slide diagram depicting oxygen diffusion that occurs during countercurrent exchange (bottom right of card; see Figure 4.1.2c for original reference diagram), yet was incorrect to include this in what was clearly a lung example, as countercurrent exchange does not occur in air-breathing organisms with lungs. While all drawn responses demonstrating CCE included some form of drawing, it is evident that students also relied heavily on verbal descriptions of the processes involved in order to convey their knowledge, as 70% of responses that were coded for CCE used a combination of words and pictures (Figure 3.1.3). Regardless of how students attempted to communicate this Key Concept, it was clear that it was a relatively difficult idea for them to fully understand in order to correctly apply it in their responses, likely due to the entry-level nature of the students taking the course and therefore the higher likelihood that they had not been exposed to this concept prior.

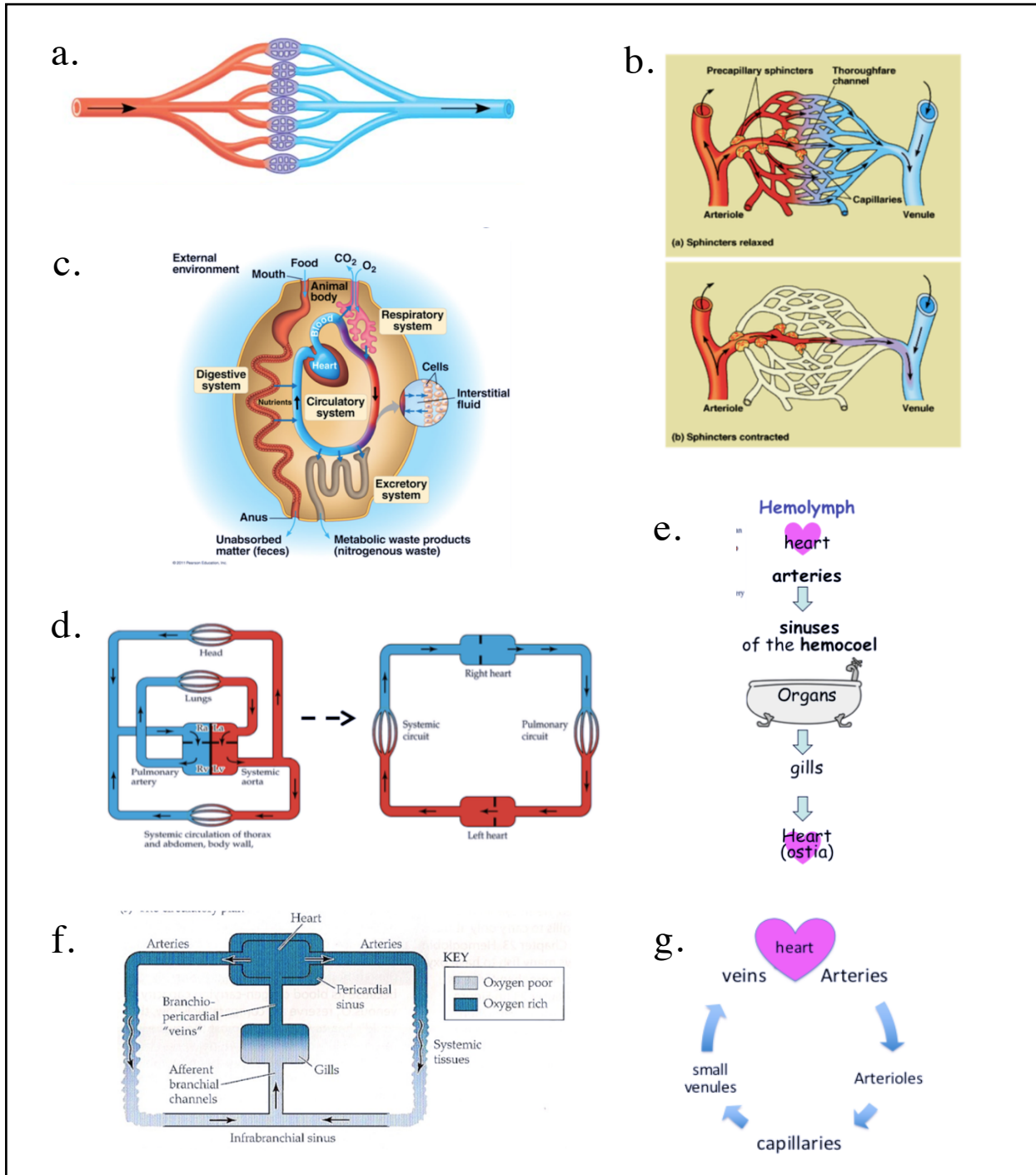


Figure 4.1.3. Examples of mass transport diagrams from lecture slides. Examples were pulled from S. Lindsay’s 2019 slides for Lectures 19 and 20. Images sources from top left: (a) from Urry, Cain, Wasserman, Minorsky, & Reece, (2015); (b) and (c) from (Reece et al., 2011); (d) and (f) Hill, Wyse, & Anderson, (2012); (e) and (g) original images by S. Lindsay.

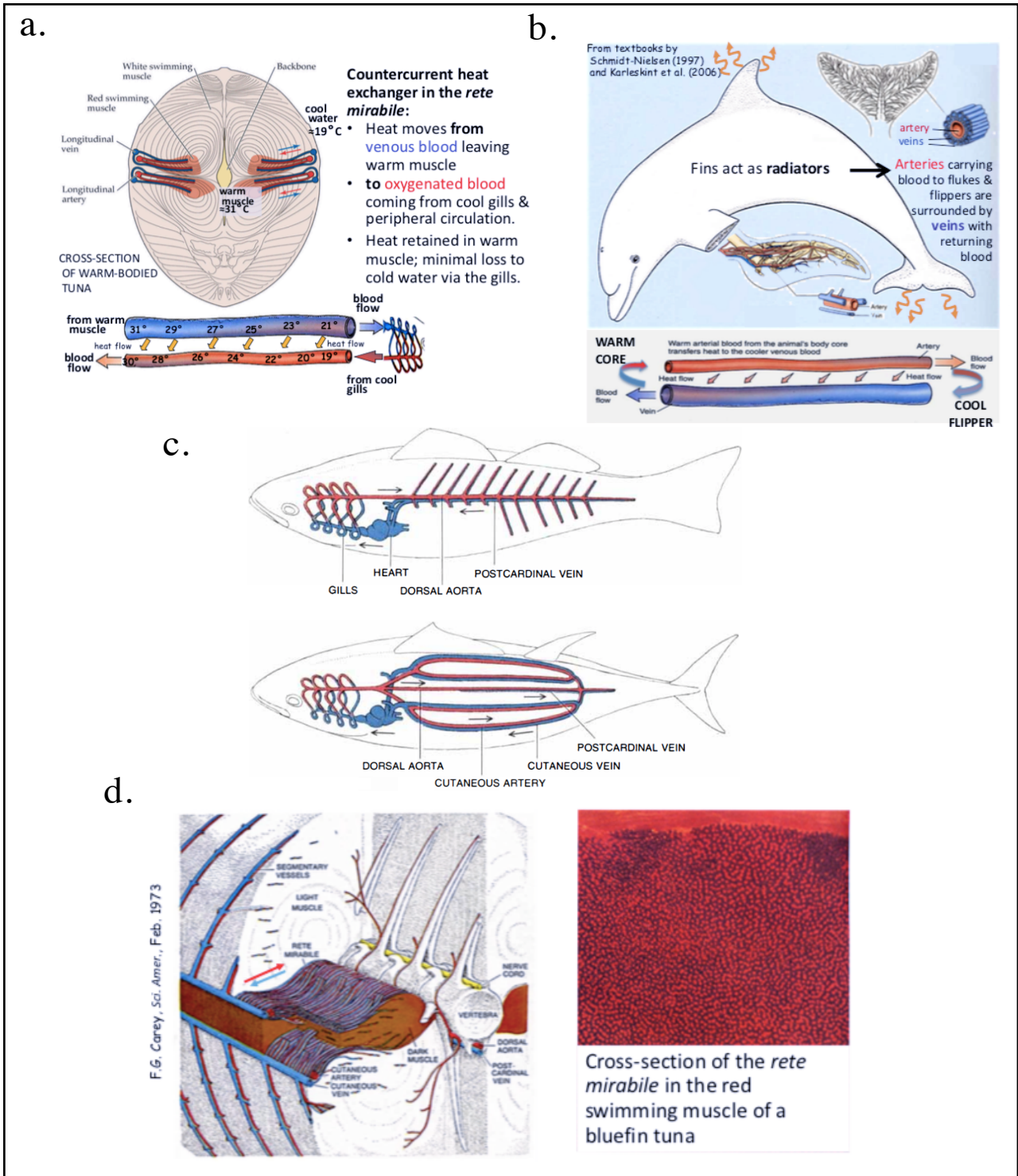


Figure 4.1.4. Examples of countercurrent exchange diagrams from lecture slides. Examples were pulled from S. Lindsay's 2019 slides for Lecture 20. Image (a) from Hill, Wyse, & Anderson, (2012); (b) from Karleskint, Small, & Turner, (2006) and Schmidt-Nielsen (1997); (c) and (d) from Carey (1973).

4.1.3 Finer-grained resolution of structures are more easily represented visually than verbally

Demonstration of student knowledge on a topic depended not only on the Key Concepts depicted in their responses, but also what representations, structures, and ideas were present; the Representation coding structure was an effort to capture these pieces. Interestingly, as part of the iterative coding process, the System-Level Depiction code was added much later than the other Representation codes, when cards were re-examined to include this code after it was anecdotally noticed that written responses appeared to neglect some of the detailed structures within each body system. Indeed, this second round of coding demonstrated that written responses more so than drawn tended to describe circulation and respiration as systems, without finer detail or discussion about the constituent structures or processes linking the two (see Figure 3.1.5).

A possible explanation for this trend could be based on how the question itself was asked, as the exact wording of the prompt was: “Describe an example to support the following statement: *Respiratory and circulatory systems* are closely linked” (emphasis added). In written responses, it is possible that students found it easier to describe their example using the wording from the prompt itself, without seeing a need to describe structures at any finer level of detail. Drawn responses, however, required students to depict their explanation visually and thus it was more difficult for responses to explicitly parrot references from the prompt itself. Given that 29.1% of responses from 2019 included the Slide Diagram code, students in the Drawn response group *did* appear to reference the abundance of slide diagrams that provided finer-grained detail of both circulatory systems and respiratory systems. This suggests that perhaps the wording of the prompt itself – and the lack of explicit prompting or option that students could support their response with a diagram or drawing – could have lead students in the Written response category

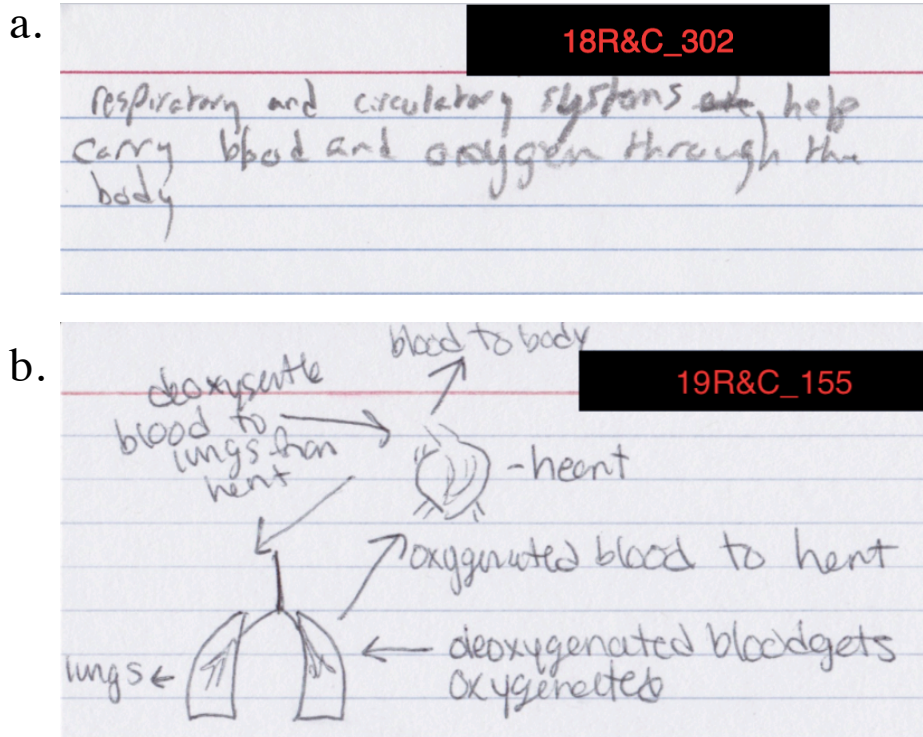


Figure 4.1.5. Illustrative student examples of “System-Level” Representation code versus finer-grained structural detail. The top card from the Written category (a) reads: “*Respiratory and circulatory systems help carry blood and oxygen through the body*” (emphasis added on phrasing mirroring the prompt wording). The bottom card from the Drawn response category (b) communicates much of the same sentiment, but expressly draws the heart and lungs.

to stick to more superficial descriptions of the biology, rather than the more nuanced and detailed approach that was more prevalent in 2019.

A striking comparison of this System-Level Depiction code versus codes for more fine-grained Representation structures can be seen in Figure 4.1.5. In comparing the actual ideas communicated on these two cards – one from the Written category (top), and one from Drawn (bottom) – they are quite similar, with both students getting at the idea that the two systems work together to move blood and oxygen throughout the body. What is particularly interesting, however, is what the 2019 student chose to draw on their card: specifically, the heart and lungs, which are finer-grained substructures of the two systems that are entirely absent from the written

response. Examples like these may indicate that students who drew found it easier to naturally weave in substructures like the heart, lungs, and/or gills to their responses in a way that was not as apparent or easily accomplished in written responses that mirrored the wording of the question itself.

4.2 Demonstration of expert thinking: drawn versus written

4.2.1 What differentiates expert thinking?

If instructors should be designing more open-ended, free-response questions to develop essential scientific skills such as visual literacy in their classrooms, it is important to ensure that they also understand what constitutes “novice” versus “expert” communication on a concept. Much research has been done on what characterizes expert learning; one emergent theme is that of information patterns and mental frameworks being highly interconnected (Committee on Developments in the Science of Learning, 2000). Knowledge organization, to experts, is comprised of informational networks that are organized around core concepts, a system that lends itself to larger “chunks” of information that can be more easily recalled, synthesized, and applied to novel situations for more effective and efficient problem-solving (Anderson & Schönborn, 2008; Committee on Developments in the Science of Learning, 2000). Novices, however, tend to see concepts within and especially across disciplines as discrete bits of information; thus, novices tend to have difficulty framing problems because of the fragmented nature of their knowledge, cannot visualize or reason through processes in an abstract way across biological scales and levels of organization, and often cannot approach problem-solving with a systems-based thinking (Anderson & Schönborn, 2008). In order to improve student problem-solving and encourage skills that skew towards expert abilities, instructors (especially in introductory level courses, but throughout all levels) therefore have to organize curricula and

assessment to reflect and support conceptual, “big picture” thinking, rather than focusing too narrowly on finer details that can be covered more thoroughly in more specialized courses (American Association for the Advancement of Science, 2011; Committee on Developments in the Science of Learning, 2000). Certain pedagogical tools, such as the BioCore instrument that helps instructors navigate and implement broader recommendations laid out in *Vision and Change*, can provide specific ways in which instructors can work towards achieving these goals in their classrooms (Brownell, Freeman, Wenderoth, & Crowe, 2014).

4.2.2 Drawn responses

Although drawing and diagramming potentially facilitated students’ abilities to communicate certain Key Concepts in response to the question prompt (e.g. mass transport and the integration of circulatory and respiratory systems), the facility of knowledge demonstration in one modality over the other does not necessarily translate to more expert-like thinking (see Section 3.2.3). Since demonstrating expertise in a subject goes beyond mere statement of facts or concepts and includes across-scale organization and thinking, as well as indicating the interconnectedness of these concepts, Key Concepts cannot stand alone to demonstrate expertise. However, it does appear that for this particular content matter, drawing and diagramming could potentially get students closer to reaching the expert bar than writing alone.

Representation code data revealed that students who drew their answers uncovered more fine-grained structural and procedural details of circulatory and respiratory systems than those who merely wrote their answers (Figure 3.1.5). The attention to substructures such as the heart, lungs or gills, and capillaries could be indicative of students starting to think across scale on an organismal level, particularly when combined with any or all of the Key Concept codes. Furthermore, Expert Score (ES) data indicated a greater percentage of written responses

receiving an ES of 0 (“prompt not addressed”) and a correspondingly higher percentage of drawn cards that received an ES 2 (“medium”) over written cards (Figure 3.2.1). Although only one response from each category was scored at an ES 3 level, this is not altogether unexpected as the course was an introductory-level class that was comprised primarily of freshmen and sophomores. Given the preceding discussion of expert-like thinking, an ES of 2 for the students in question is a laudable achievement in its own right, and the greater percentage of students from the 2019 cohort achieving an ES of 2 is therefore noteworthy. Finally, the inclusion of more Key Concepts in drawn responses over written (Figure 3.1.1) can also help explain higher Expert Scores achieved in the Drawn category. The nature of the Key Concepts themselves including systems-level concepts (e.g. with mass transport), as well as reasoning across scales: countercurrent exchange occurs on a molecular level, circulatory-respiratory integration on an organ level, and mass transport on a whole-organism level. Thus, inclusion of more than one of these Key Concepts on a card by a student tends to lend itself to communication that is more aligned with expert-like discourse.

Taken together, these data suggest that explicitly prompting students to draw their responses gets students thinking in a way that encourages a more expert-like approach to tackling the question posed and reduces the risk that students will miss the mark entirely in their answers. It is possible that for students who struggled to formulate a response to the prompt, *any* kind of drawing conveys more and/or better information than a random attempt at writing their response. Furthermore, although it was beyond the scope of this study, it may be useful to consider integrating drawn responses into classrooms with English as a Second Language (ESL) learners. Giving ESL students the opportunity to demonstrate their knowledge in drawn form could potentially remove language and vocabulary barriers that would otherwise impede their ability to

express themselves, and could provide instructors with useful feedback on student knowledge that could otherwise be inaccessible. Indeed, drawing-to-learn may be a means by which instructors can address equity issues in the classroom in which narrow assessment types unintentionally exclude or unfairly punish ESL and other nontraditional learners, as well as limit instructor knowledge on such students' capabilities (Darling-Hammond, 1994).

4.3 Alternative Conceptions and their impact on determining expertise

Alternative conceptions, or preconceptions, are inaccurate ideas or mischaracterizations of phenomena that occlude students' thinking and understanding; however, it appears that not all alternative conceptions are created equally. Arguably, certain alternative conceptions are less detrimental to students achieving expert-like thinking than others. For instance, students may depict a multitude of factual ideas on their responses that communicate deeper content knowledge and thinking across scales, but a certain sentence or diagram may warrant an Unclear Intent code; this would not necessarily invalidate the rest of the notecard. Conversely, a card that states or depicts a clear comparison between the respiratory and circulatory systems, such as "they are closely linked because they both contain structures that branch," demonstrates a lack of understanding of the question and a more surface-level, novice approach to the content that warrants a lower assessment of the student's expertise. In assessing Alternative Conception (AC) codes that appeared across novice responses and response types, Unclear Intent and Terminology were the most frequently observed ACs. Since these codes account for students demonstrating or depicting vague, incomplete, or inaccurate information, it is not surprising to see that novice responses (ES of 0 or 1) would have high frequencies of these ACs out of all AC codes tagged. However, other AC codes appear less straightforward in the trends observed, and appear to

reveal different levels of thinking that are brought forth by writing versus drawing a response, warranting a closer consideration of the best means of asking a question.

4.3.1 Certain alternative conceptions lend themselves to appearing in written form

A few of the Alternative Conception codes in this study – Terminology, Comparison, and “Vitalistic” Response – appeared more frequently in written than drawn responses; see Figure 3.3.1. Terminology is perhaps the least surprising of the three, as this AC code referred to either a misuse of terms, confusing descriptions, and/or vague written explanations of the processes depicted on the notecard, all of which one would expect to see at higher frequencies in strictly textual responses versus visual. Much like the Terminology code, the Comparison AC code, while present in both written and drawn, was often cued by words rather than visuals. For example, “both,” “between,” or “similar” were key words that students used to compare the two systems, which were more common to see in 2017 and 2018 responses than in 2019 due to the nature of response types.

Along these same lines, the “Vitalistic” Response AC code would be difficult to demonstrate in a drawing. Cards that included this code had phrasing such as: “...blood is needed for the respiratory system to function properly,” “...blood is needed for respiration...,” “...both [are] vital for life,” “If you stop breathing, your heart stops receiving oxygen and all the blood it pumps will be anoxic,” and “You need to breath[e] to keep your heart pumping,” (see Figure 4.3.1). These “vital sign” responses linking the two systems are not necessarily inaccurate, but do not convey the level of detail and expertise required to demonstrate deeper thinking about the content matter. Further, these are difficult concepts to convey in drawn form; how does one draw the need for oxygen to keep a heart pumping? The higher frequencies of all

three of these AC codes in written versus drawn responses therefore are not altogether unexpected, given the specific prompt in 2019 for students to draw their responses.

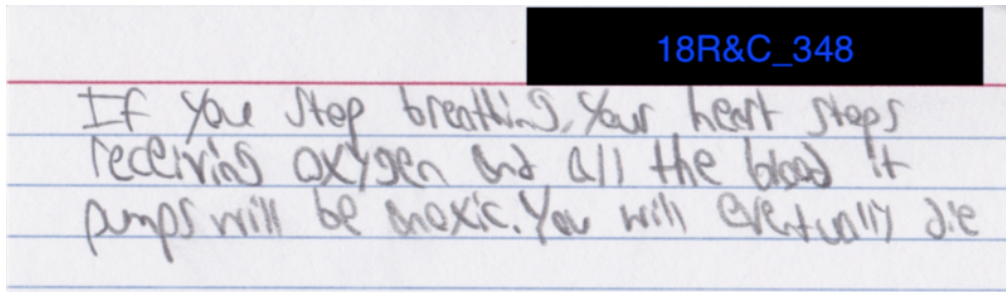


Figure 4.3.1. Written example of “Vitalistic” Response Alternative Conception code. Response reads: “If you stop breathing, your heart stops receiving oxygen and all the blood it pumps will be anoxic. You will eventually die.”

4.3.2 Earning the “Capillaries Absent” Alternative Conception code itself requires a level of expertise

One Alternative Conception code that was significantly higher in drawn responses over written was that of Capillaries Absent, which requires a deeper analysis to understand the factors affecting a card “earning” this particular code. Capillaries Absent indicates that students were missing a key piece of information; this code was a child code of the Key Component Absent parent code within Alternative Conceptions, described as “Drawing or explanation lacks critical component of system for explanation/ response to be considered complete” (Appendix A, Key Component Absent). The “decision tree” to get to this code, therefore, required that students at a minimum communicate the concepts of structures that contain the capillaries, which could take the shape of a circuit diagram that included lungs, gills, and/or body tissues, or a description thereof, and that capillaries were then expressly omitted from the response (e.g. Figure 4.3.2). Correspondingly, students who described or drew these circuits or structures already displayed a certain command over the subject matter, but merely fell short of a truly expert-like explanation

that included descriptions of surface area-to-volume tradeoffs and/or close investment of capillaries. Thus, the trends observed and described from Figure 3.3.2 are not unexpected; it is *expected* that Capillaries Absent would be seen at higher frequencies in higher Expert Score (ES) categories, as truly novice responses (ES 0, and to some extent ES 1 responses) would not even meet the criteria to be considered for the Key Component Absent decision tree.



Figure 4.3.2. Drawn example of Capillaries Absent Alternative Conception code. In this response from 2019, the student has depicted the gills, heart, and the branching nature of the circulatory system accurately, as well as included the critical oxygen molecule. However, oxygen is depicted as moving directly from the gills to the heart with no mention or depiction of capillaries.

Another point to consider is why the Capillaries Absent code occurs at a significantly higher frequency in drawn versus written responses (Figure 3.3.1). This trend may be explained by considering the points made in Section 4.1, in which it was argued that drawing can lend itself to communicating both higher resolution of structures and more Key Concepts. Taken together with the claim that students must reach a higher level of content communication before the Capillaries Absent code can even be considered, it is reasonable to suggest that the drawn responses more frequently demonstrate these criteria; thus, the Capillaries Absent code is more likely to appear in drawn responses. Interestingly, this suggests that for this particular notecard question, there may be utility in prompting students to draw their responses if the drawn category

appears to be uncovering higher-order thinking in a way that written responses are not capturing. As instructors construct questions related to circulatory and respiratory in future courses, or indeed other marine science and biology courses, they may do well to consider expressly training students in drawing methods in order to include drawing prompts in formative and summative assessments.

4.4 Limitations and contributions of current study

Although this study provided great insight into students' verbal and pictorial representations of a core physiological concept, questions remain. Potential limitations resulting from the current study design and execution are discussed below that are important to keep in mind when considering the data and conclusions thus far; however, there are still valuable applications to be gleaned from the results.

4.4.1 Scaffolding drawing opportunities during instruction may remove the barrier to entry

Though drawing-to-learn strategies should not necessarily be universally applied to all marine science education, the results from this study suggest that providing drawing opportunities for students to demonstrate their knowledge on a topic can provide more targeted insight into their thinking and/or understanding of the content matter. Indeed, given the immense amount of detail that can be conveyed about the circulatory and respiratory systems of marine mammals in one simple diagram (e.g. any one of the panels from Figure 4.1.3), drawings may certainly be a better and/or more efficient tool for students to use to convey knowledge in an assessment question. One possible reason that drawn responses in this study were not universally outperforming written, however, could be that the barrier to entry for drawn responses was higher than that of written. Although there were a handful of drawing-to-learn opportunities

presented as in-class activities and homework in 2019 (including a diagramming of worm body plans and a concept map for primary production), many students expressed discomfort during these activities with being able to draw fluently so as to convey information, particularly during the worm body plan activity (Siddons, personal observation). It would seem that many students struggle with making a distinction between being able to draw to produce realistic and artistic visuals, versus using drawing – however simplified or abstract – as another tool in their academic toolkit.

Employing visual teaching strategies can have the cognitive benefit of reducing learner cognitive load by providing mental scaffolds for students to insert learned material (Gross et al., 2017), explicitly linking text to diagrams (Mason et al., 2013), and reducing extraneous information to highlight key information (Mayer, 2018). Cognitive load is the amount of information a learner must hold in a mental framework at any one time; a variety of techniques exist to reduce the amount of mental processing a lesson demands of the learner through a “less is more” approach in order to maximize educational output of an activity (Mayer & Fiorella, 2014). Alternatively, providing students engaged in a learner-generated drawing activity with standardized representative symbols, a word bank, or the outline of a drawing to be filled in can also help reduce the cognitive load that would have otherwise been required for a student to generate one or more of these details his or herself. This allows the learner to focus on the more relevant and educationally important task of learning the key information to be captured in the drawing itself (Mayer & Moreno, 2010).

In this study, students were likely somewhat supported in reducing cognitive load when they were asked to draw by the heavily visual nature of the slides and lecture materials themselves. A 2017 study by Gross, Wright, and Anderson used an image-based activity in an

undergraduate anatomy course resulted in higher exam scores than those students who only participated in text-based learning activities (Gross et al., 2017). Based on quantitative and qualitative survey data, the authors concluded that image-based learning provided students with a mental scaffold with which to organize course material, thereby reducing cognitive load and focusing attention on the relevant information to be learned (Gross et al., 2017). Similarly, students in this introductory course were likely leaning on images they saw on slides – and perhaps copied into their notes to some degree for later reference – to both facilitate their own SOI mental process, as well as direct how draw or diagram their responses in 2019. However, students may have been stymied by how to depict certain elements in drawn form that didn't directly correlate to images that they had already seen in the slide diagrams.

Another study (Mautone & Mayer, 2007) specifically sought to understand how scaffolding supported the organization and integration cognitive steps in the SOI model of learning. In this example, the researchers provided students with different types of scaffolding supports to aid in graph interpretation: signaling and structural graphic organizers provided scaffolds for the organization of material, whereas concrete graphic organizers provided integrative support. Student-generated statements responding to prompts and questions about the graphs were then classified as relational, descriptive statements that reflected correct organization of the material, or causal statements that reflected correct integration (Mautone & Mayer, 2007). Students who were provided with organizational or integrative scaffolds produced more correct statements of the corresponding type; additionally, students provided with *any* type of scaffold versus those who were not demonstrated deeper learning of the material, suggesting that strategic implementation of scaffolding can increase learner engagement and interpretation of data presented in graphs, and reduce the effort expended on trying to understand basic

graphical features such as axes and legends (Mautone & Mayer, 2007). Because there was not explicit instruction or supports such as these employed in this study, students may have struggled with interpreting some of the more complicated diagrams – such as those conveying countercurrent exchange – such that they were unable to later communicate those concepts in a meaningful or accurate way in their responses.

To help remedy this, future studies may want to scaffold drawing-to-learn more rigorously so students are as comfortable using drawing to communicate ideas as they are with the written word. Providing more scaffolded drawing activities in class, such as giving students “symbol banks” to use in their drawings, much like a word bank is provided for written responses, could help students feel less intimidated about approaching how to represent certain structures or processes visually. With more practice and training in this skill, a repeat study that trains all students in drawing techniques, but specifically instructs half the group to write and half to draw their responses to the prompt could help make a clearer case for crafting assessment questions to explicit instruct students to respond one way or the other. Of course, researchers would need to consider whether the subject and/or content that is being assessed lends itself to potential drawing or diagramming before embarking on such a study.

4.4.2 Assessment type may affect student effort in response

In this study, the notecard questions posed were collected as a means of quick formative assessment to gauge student knowledge on a particular topic covered in lecture thus far, counting only for participation credit in the class. A major advantage to using this type of formative assessment tool is that it addresses some of the previously cited issues faculty may have with using formative assessments in large lecture-style undergraduate courses (Goubeaud, 2010); namely, that there are time constraints and class sizes are too large to effectively implement

formative assessment opportunities. Although students had a required amount of participation credit built into their overall course score, and the notecard questions were unannounced in order to promote attendance at lecture, it is possible that in the context of overall course grades, the notecard questions provided little incentive for students to expend top effort on their responses. Assessing student responses on more high-stakes formats such as particular questions on summative exams may therefore provide more insight into whether there is indeed an effect on student effort based on their perception of stakes and/or incentives to perform well. A comparison of in-class formative assessment data like the notecards, alongside matched questions on summative assessments that target the same content and give students more time to think about their responses, could also potentially capture a better representation of what students actually know than the data collected in the present study.

4.4.3 Notecards alone cannot capture the total picture of student understanding

Perhaps the most important takeaway from the current study design is that the notecards themselves are a snapshot of student knowledge in that moment of time, and could in fact be better thought of as snapshots of what the students chose to *communicate* as their knowledge, rather than a complete picture of their actual knowledge of content. In addition to potentially lacking incentivization for students to try their best, the timing and implementation of the lecture notecard format – given at the end of lecture with only a few minutes to complete – could further complicate whether the data is truly representative of the depth and/or breadth of student knowledge. It is important to keep in mind that the present study does not presume to make conclusions about any given student or group of students' total *understanding* of the subject matter. Future studies, therefore, may want to consider supplementing notecard data with interviews with individual students to ask participants to describe what they are thinking when

writing versus drawing their responses. This could provide critical insight into whether student understanding from interview data aligns with the inferences drawn from the notecard data, particularly in comparing interviews with overall notecard Expert Scores. Indeed, students who have a deep understanding of the content but chose – either due to time constraints or lack of incentive – to just write or draw a quick response and pass in their notecard could be miscategorized as novices, when in fact their thinking is more aligned with expert-like tendencies (or vice versa).

4.4.4 Contributions to marine science education

There are several key takeaways from this study that bear mentioning, as they contribute to the research in marine science education and drawing-to-learn strategies. First, this study adds to the discourse on how questions or assessments are posed in coursework, and encourages marine science instructors to think beyond traditional selected-response questions like multiple choice that may constrain student expression of knowledge or alternative conceptions. By implementing diagramming and drawing constructed-response questions strategically to concepts that are visual-spatial in nature and ask students to make cross-cutting connections between concepts that may be otherwise taught in isolation, the current study illustrates one answer to the direct call for tested assessments from studies such as Weatherbee and Lindsay (2018). Indeed, the 2018 study specifically cited students' difficulty with graph interpretation and application (Weatherbee & Lindsay, 2018), a form of visual representation that is critical to scientific discourse not only in marine science, but across all STEM disciplines. By having students engage in creating their own visual representations of information and data in a formative setting, this study identifies an avenue by which instructors can address some of these interpretive shortcomings.

The methodological approach used here can also be modified to help biology and marine science instructors identify key ideas, structures, and alternative conceptions in visual question prompts used in their own classrooms. As instructors revisit question prompts and topics to determine if they are suited for visual representation and reasoning, the coding structure used here can be used as a guideline for assessing student-produced diagrams for quality and accuracy of response. Instructors can create quick lists of their own key concepts, representations, and common alternative conceptions – three key pieces from student responses that informed the research conclusions here – ahead of time based on their own knowledge and experience to determine what a coherent and accurate student response (drawing or diagram) should contain. While it is not reasonable to expect instructors to apply the full coding schematic described here to every student response for every question prompt posed in a class, particularly given the size of many university courses, the three key pieces mentioned can serve as a rubric that both instructors and students can use: instructors as an assessment tool, and students as a means to inform, guide, and correct their own thinking on a topic. In doing so, instructors can also create a space in class or online for more dialogue on the diagrams themselves in an effort to address visual literacy goals.

Finally, incorporating more assessment prompts that ask students to create a visual constructed-response such as the one described in this study can address larger instructional goals outlined in *Vision and Change* (AAAS, 2011). As mentioned in Quillin and Thomas (2015), one of the core competencies this document calls for is “Modeling and Simulation,” yet it is only described in the *mathematical* sense; using diagrams or drawings to model or simulate is entirely absent in the description of how this competency could look like in the classroom. In implementing drawing-to-learn strategies, instructors should consider whether the instructional

goal is the *product* (as was assessed in this study), the *process*, or both. The means by which the assessment prompt presented here was delivered could be modified to include some of the more explicit conversations about diagrammatic choices and conventions in order to address this *process* piece. More instructional time spent on analyzing graphs, diagrams, and drawings for their meaning as well as specific graphic choices (e.g. conventions such as arrows, color choice, etc.) can help address and correct specific misunderstandings or alternative conceptions that students hold (e.g. Evagorou et al., 2015) – such as the common misunderstanding that deoxygenated blood is actually blue, due to a combination of how blood vessels appear under the skin as well as the diagrammatic convention of showing deoxygenated blood as blue and oxygenated blood as red in diagrams (e.g. Figure 4.1.2a). Asking students to make explicit their reasoning behind visual representation choices can also illuminate and uncover the selection – organization – integration mental framework, which can help teachers tailor instruction accordingly. Additionally, such conversations can shift the focus on using diagrams and drawings as models *of* a process or concept, to using them as models *for* understanding.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

This study demonstrated that for certain marine science topics, particularly those related to organismal biology, more opportunities for students to practice drawing-to-learn activities could prove beneficial to both learning and for providing evidence of content knowledge. Prompting students to draw and diagram their responses, as opposed to writing alone, allowed students to communicate more key ideas, as well as more detailed responses of organism structure and function. The sample of students who drew/diagrammed their responses achieved higher Expert Scores for their responses than those in the written years. Although alternative conceptions appeared in both types of student responses, the *type* of alternative conception varied between written and drawn categories.

Given the findings, a few key recommendations are outlined for instructors wishing to explore drawing-to-learn opportunities in their own classrooms:

1. *Explicitly teach students how to navigate and critique drawings and diagrams present in expository texts.*

Many instructors assume that students already know how to read and interpret diagrams and drawings that appear in expository texts that are used in classroom instruction. Include explicit instruction in lectures for how to read different visual representations of information, particularly field- or content-specific representations that are conventionally understood by experts (e.g. \downarrow representing the suppression of gene expression by geneticists), but may be foreign to students. Foster conversations with students about how choices in color, size, shape, and other visual aspects can impact the message being communicated in diagrams, and encourage opportunities for students to critically evaluate diagrams in their course materials.

2. *Scaffold drawing-to-learn activities and provide plenty of opportunities for practice*

Along with teaching students how to be critical visual consumers, scaffolding how to use drawing to learn and communicate content is crucial if students are expected to use drawing on assessments. Much like the concept of providing a “word bank” for students to use when writing a response, “symbol banks” may be employed to help lower the barrier for entry for students who find themselves stuck when asked to draw (e.g. telling students to use circles to represent K^- ions and triangles to represent Na^+ ions in a diagram of membrane potential). Other strategies may include providing more difficult parts of the diagram to draw, such as the lipid bilayer for membranes, and asking students to fill in simpler parts of the diagram. Additionally, modeling how to diagram and providing students with examples of diagrams that are more abstract than representational could help students who are intimidated by drawing better utilize drawing-to-learn strategies. Most of all, provide students with plenty of in-class and out-of-class opportunities to practice drawing and diagramming prior to requiring any such efforts on high-stakes summative assessments such as preliminary or final exams.

3. *Implement low-stakes formative assessments on a topic-by-topic basis with drawing-to-learn opportunities to “test drive” if the content lends itself to drawing before implementing on higher-stakes formats like exams or other summative assessments.*

Not all disciplines, or content matter within disciplines, can be easily depicted in drawn form, and may actually be best described verbally. Prior to asking students to draw responses on exam questions, offering low-stakes formative opportunities (like lecture notecard questions) can give instructors the opportunity to “test drive” if drawing-to-learn is a useful tool to employ in a given subject or topic.

4. *Give students the opportunity to write OR draw in their responses to exam questions.*

Finally, when certain subjects or content areas are determined to be conducive to drawing or diagramming, allowing students the option to write, draw, or both when responding may expand opportunities for them to demonstrate knowledge on a topic and give instructors a clearer picture of their understanding.

Future research can help develop a fuller understanding of the different aspects at play when students are asked to draw and diagram in the classroom. A more robust data collection, including analysis of student notebooks during the course of instruction and assessment, combined with selected student interviews on both notebook content and notecard response, could better clarify individual students' SOI processes and give insight into what diagrams or drawings were most useful to student learning. Additionally, such data could help address what underlies student motivation to draw (or not draw) in their assessment responses, and if there is any influence on self-efficacy in drawing-to-learn. For instance, future research questions may ask how students perceive their own drawing skills, and if that impacts how they utilize drawing-to-learn in traditionally "non-artistic" spaces such as STEM classrooms. Finally, research into the application of the drawing-to-learn strategies used in this study could be expanded to other topics within marine science, including non-biological subjects such as physical oceanography, to determine other content areas in which student learning could benefit from drawing and diagramming.

Drawing-to-learn can provide a rich learning experience for both students and instructors alike, as there are opportunities for more meaningful learning and better communication of content. Further, developing fluency in visualizations and diagrammatic representations give students the opportunities to communicate ideas and data as disciplinary experts do, and

moreover provides more authentic interaction with the course material. Given the recommendations and aims set forth in the *Vision and Change* document, drawing-to-learn in an undergraduate setting may provide a key avenue for university faculty to implement these goals in their own coursework.

REFERENCES

- ACRL. (2012). Visual literacy competency standards for higher education. *College & Research Libraries News*, 73(2), 97–104. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=71495825&lang=tr&site=ehost-live>
- Ainsworth, S., Prain, V., & Tytler, R. (2011). Drawing to Learn in Science. *Source: Science, New Series*, 333(6046), 1096–1097. <https://doi.org/10.1111/j.1756-8765.2011.01149.x>
- Airey, J., & Linder, C. (2009). A disciplinary discourse perspective on university science learning: Achieving fluency in a critical constellation of modes. *Journal of Research in Science Teaching*, 46(1), 27–49. <https://doi.org/10.1002/tea.20265>
- American Association for the Advancement of Science. (2011). *Vision and Change in Undergraduate Biology Education: A Call to Action*. Washington, DC. Retrieved from www.sciencetranslationalmedicine.org
- Anderson, T. R., & Schönborn, K. J. (2008). Bridging the educational research-teaching practice gap. Conceptual understanding, part 1: The multifaceted nature of expert knowledge. *Biochemistry and Molecular Biology Education*, 36(4), 309–315. <https://doi.org/10.1002/bmb.20209>
- Arthurs, L., Hsia, J. F., & Schweinle, W. (2015). The Oceanography Concept Inventory: A Semicustomizable Assessment for Measuring Student Understanding of Oceanography. *Journal of Geoscience Education*, 63(4), 310–322. <https://doi.org/10.5408/14-061.1>
- Backhouse, M., Fitzpatrick, M., Hutchinson, J., Thandi, C. S., & Keenan, I. D. (2017). Improvements in anatomy knowledge when utilizing a novel cyclical “Observe-Reflect-Draw-Edit-Repeat” learning process. *Anatomical Sciences Education*, 10(1), 7–22. <https://doi.org/10.1002/ase.1616>
- Barrett, B. S., Swick, W. A., & Smith Jr., D. E. (2014). Assessing an undergraduate oceanography curriculum. *Oceanography*, 27(4), 13–17. <https://doi.org/10.5670/oceanog.2014.99>
- Beatty, N. A. (2013). Cognitive Visual Literacy: From Theories and Competencies to Pedagogy. *Art Documentation: Journal of the Art Libraries Society of North America*, 32(1), 33–42. Retrieved from <http://www.journals.uchicago.edu/t-and-c>
- Brody, M. J., & Koch, H. (1990). An assessment of 4th-, 8th-, and 11th-grade students’ knowledge related to marine science and natural resource issues. *Journal of Environmental Education*, 27(3), 16–26. <https://doi.org/10.1080/00958964.1990.9941927>

- Brownell, S. E., Freeman, S., Wenderoth, M. P., & Crowe, A. J. (2014). BioCore Guide: A Tool for Interpreting the Core Concepts of Vision and Change for Biology Majors. *CBE- Life Sciences Education*, 13, 200–211. <https://doi.org/10.1187/cbe.13-12-0233>
- Carey, F. G. (1973). Fishes with warm bodies. *Scientific American*, 228(2), 36–44.
- Clark, J. M., & Paivio, A. (1991). Dual Coding Theory and Education. *Educational Psychology Review*, 3(3), 149–210. Retrieved from <https://www.jstor.org/stable/23359208>
- Committee on Developments in the Science of Learning. (2000). How experts differ from novices. In *How People Learn: Brain, Mind, Experience, and School, Expanded Edition* (pp. 31–50). National Academies Press. <https://doi.org/10.1136/gutjnl-2011-301145>
- Corbin, J. (2013). Strauss' grounded theory. In C. T. Beck (Ed.), *Routledge International Handbook of Qualitative Nursing Research* (pp. 169–182). New York: Routledge.
- Couch, B. A., Hubbard, J. K., & Brassil, C. E. (2018). Multiple-True-False Questions Reveal the Limits of the Multiple-Choice Format for Detecting Students with Incomplete Understandings. *BioScience*, 68(6), 455–463. <https://doi.org/10.1093/biosci/biy037>
- Cudaback, C. (2006). What do college students know about the ocean? *Eos*, 87(40). <https://doi.org/10.1029/2006EO400003>
- Cudaback, C. (2008). Ocean Literacy: There's More to it Than Content. *Oceanography*, 21(4), 10–11. <https://doi.org/10.1029/2006EO400003>
- Cummins, S., & Snively, G. (2000). The Effect of Instruction on Children's Knowledge of Marine Ecology, Attitudes Toward the Ocean, and Stances Toward Marine Resource Issues Shirley Cummins & Gloria Snively, University of Victoria, Canada. *Canadian Journal of Environmental Education*, 5, 306–324.
- Darling- Hammond, L. (1994). Performance-Based Assessment and Educational Equity. *Harvard Educational Review*, 64(1), 5–30.
- Doucette, J. (n.d.). The Ocean Conveyor. Woods Hole Oceanographic Institute. Retrieved from <https://www.whoi.edu/know-your-ocean/ocean-topics/ocean-circulation/the-ocean-conveyor/>
- Edens, K. M., & Potter, E. (2003). Using Descriptive Drawings as a Conceptual Change Strategy in Elementary Science. *School Science and Mathematics*, 103(3), 135–144.
- Eilertsen, T. V., & Valdermo, O. (2000). Open-book assessment: A contribution to improved learning? *Studies in Educational Evaluation*, 26, 91–103. Retrieved from www.elsevier.nl/stueduc

- Evagorou, M., Erduran, S., & Mäntylä, T. (2015). The role of visual representations in scientific practices: from conceptual understanding and knowledge generation to ‘seeing’ how science works. *International Journal of STEM Education*, 2(1), 1–14. <https://doi.org/10.1186/s40594-015-0024-x>
- Fan, J. E. (2015). Drawing to Learn: How Producing Graphical Representations Enhances Scientific Thinking. *Translational Issues in Psychological Science*, 1(2), 170–181. <https://doi.org/10.1037/tps0000037>
- Felten, P. (2010). Visual Literacy. *Change: The Magazine of Higher Learning*, 40(6), 60–64. <https://doi.org/10.3200/CHNG.40.6.60-64>
- Fiorella, L., & Mayer, R. E. (2016). Eight Ways to Promote Generative Learning. *Educational Psychology Review*, 28, 717–741. <https://doi.org/10.1007/s10648-015-9348-9>
- Fortner, R. W., & Mayer, V. J. (1989). Marine and aquatic education—a challenge for science educators. *Science Education*, 73(2), 135–154. <https://doi.org/10.1002/sce.3730730203>
- Gilbert, J. K. (2005). Visualization: A Metacognitive Skill in Science and Science Education. In J. K. Gilbert (Ed.), *Visualization in Science Education* (pp. 9–27). Dordrecht: Springer.
- Goubeaud, K. (2010). How is Science Learning Assessed at the Postsecondary Level? Assessment and Grading Practices in College Biology, Chemistry and Physics. *Journal of Science Education and Technology*, 19, 237–245. <https://doi.org/10.1007/s10956-009-9196-9>
- Gough, A. (2017). Educating for the marine environment: Challenges for schools and scientists. *Marine Pollution Bulletin*, 124, 633–638. <https://doi.org/10.1016/j.marpolbul.2017.06.069>
- Gross, M. M., Wright, M. C., & Anderson, O. S. (2017). Effects of image-based and text-based active learning exercises on student examination performance in a musculoskeletal anatomy course. *Anatomical Sciences Education*, 10(5), 444–455. <https://doi.org/10.1002/ase.1684>
- Guest, H., Lotze, H. K., & Wallace, D. (2015). Youth and the sea: Ocean literacy in Nova Scotia, Canada. *Marine Policy*, 58, 98–107. <https://doi.org/10.1016/j.marpol.2015.04.007>
- Heideman, P. D., Flores, K. A., Sevier, L. M., & Trouton, K. E. (2017). Effectiveness and adoption of a drawing-to-learn study tool for recall and problem solving: Minute sketches with folded lists. *CBE Life Sciences Education*, 16(2), 1–13. <https://doi.org/10.1187/cbe.16-03-0116>
- Hess, K. K. (2006). *Cognitive complexity: Applying Webb DOK Levels to Bloom’s Taxonomy*. Dover, NH. Retrieved from http://www.nciea.org/publications/DOK_ApplyingWebb_KH08.pdf

- Hill, R. W., Wyse, G. A., & Anderson, M. (Eds.). (2012). *Animal Physiology* (3rd ed.). Sunderland: Sinauer Associates, Inc.
- Hubbard, J. K., Potts, M. A., & Couch, B. A. (2017). How Question Types Reveal Student Thinking: An Experimental Comparison of Multiple-True-False and Free-Response Formats. *CBE- Life Science Education*, *16*, 1–13. <https://doi.org/10.1187/cbe.16-12-0339>
- Johnson, L. (2014). Adapting and combining constructivist grounded theory and discourse analysis: A practical guide for research. *International Journal of Multiple Research Approaches*, *8*(1), 100–116. <https://doi.org/10.5172/mra.2014.8.1.100>
- Karleskint, G., Small, J., & Turner, R. (2006). *Introduction to Marine Biology* (2nd ed.). Pacific Grove: Brooks/Cole Publishing Company.
- Lambert, J. (2005). Students' Conceptual Understandings of Science After Participating in a High School Marine Science Course. *Journal of Geoscience Education*, *53*(5), 531–539. <https://doi.org/10.5408/1089-9995-53.5.531>
- Lambert, J. (2006). High School Marine Science and Scientific Literacy: The promise of an integrated science course. *International Journal of Science Education*, *28*(6), 633–654. <https://doi.org/10.1080/09500690500339795>
- Lemke, J. L. (2015). Multimedia and discourse analysis. *The Routledge Handbook of Discourse Analysis*, 79–89. <https://doi.org/10.4324/9780203809068.ch6>
- Leutner, D., & Schmek, A. (2014). The Generative Drawing Principle in Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (2nd ed., pp. 433–448). New York: Cambridge University Press.
- Lucrezi, S., Milanese, M., Danovaro, R., & Cerrano, C. (2018). “Generation Nemo”: motivations, satisfaction and career goals of marine biology students. *Journal of Biological Education*, *52*(4), 391–405. <https://doi.org/10.1080/00219266.2017.1385509>
- Mason, L., Pluchino, P., & Tornatora, M. C. (2013). Effects of Picture Labeling on Science Text Processing and Learning: Evidence From Eye. *Source: Reading Research Quarterly*, *48*(2), 199–214. <https://doi.org/10.1002/rrq.41>
- Mautone, P. D., & Mayer, R. E. (2007). Cognitive Aids for Guiding Graph Comprehension. *Journal of Educational Psychology*, *99*(3), 640–652. <https://doi.org/10.1037/0022-0663.99.3.640>
- Mayer, R. E. (2014). Cognitive Theory of Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (2nd ed., pp. 43–71). New York: Cambridge University Press.

- Mayer, R. E. (2018). Designing multimedia instruction in anatomy: An evidence-based approach. *Clinical Anatomy*, 9999(August), 1–10. <https://doi.org/10.1002/ca.23265>
- Mayer, R. E., & Fiorella, L. (2014). Principles for Reducing Extraneous Processing in Multimedia Learning: Coherence, Signaling, Redundancy, Spatial Contiguity, and Temporal Contiguity Principles. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (2nd ed., pp. 279–315). New York: Cambridge University Press.
- Mayer, R. E., & Moreno, R. (2010). Nine Ways to Reduce Cognitive Load in Multimedia Learning. *Educational Psychologist*, 38(1), 43–52. https://doi.org/10.1207/S15326985EP3801_6
- McCrudden, M. T., & Rapp, D. N. (2017). How Visual Displays Affect Cognitive Processing. *Educ Psychol Rev*, 29, 623–639. <https://doi.org/10.1007/s10648-015-9342-2>
- McTigue, E. M., & Flowers, A. C. (2011). Science Visual Literacy: Learners' Perceptions and Knowledge of Diagrams. *The Reading Teacher*, 64(8), 578–589. Retrieved from <https://www-jstor-org.prxy4.ursus.maine.edu/stable/pdf/41203457.pdf?refreqid=excelsior%3Ad03bc9de17c7e45b514130aa2e5dee58>
- Mey, G., & Dietrich, M. (2017). From Text to Image- Shaping a Visual Grounded Theory Methodology. *Historical Social Research*, 42(4), 280–300.
- Molnar, C., & Gair, J. (2015). Circulatory and Respiratory Systems. In *Concepts of Biology: 1st Canadian Edition*. Victoria, B.C.: BCcampus. Retrieved from <https://opentextbc.ca/biology/>
- Momsen, J., Offerdahl, E., Kryjevskaiia, M., Montplaisir, L., Anderson, E., & Grosz, N. (2013). Using Assessments to Investigate and Compare the Nature of Learning in Undergraduate Science Courses. *Life Sciences Education*, 12, 239–249. <https://doi.org/10.1187/cbe.12-08-0130>
- Moore, C., Moore, M., Trumble, S., Niemeyer, M., Lentell, B., McLellan, W., ... Fahlman, A. (2014). A comparative analysis of marine mammal tracheas. *The Journal of Experimental Biology*, 217, 1154–1166. <https://doi.org/10.1242/jeb.093146>
- NGSS. (2013). *Next Generation Science Standards: For States, By States*. Washington, DC.
- NMEA. (2010). *Special Report #3: Ocean Literacy Scope and Sequence for Grades K-12*. Retrieved from www.marine-ed.org
- Ocean Literacy Network. (2013). *Ocean Literacy: The Essential Principles and Fundamental Concepts of Ocean Sciences for Learners of All Ages*. Retrieved from <https://www.coexploration.org/oceanliteracy/documents/OceanLitChart.pdf>

- Ocean Literacy Network. (2015). Next Generation Science Standards Aligned to Ocean literacy. Retrieved from <http://oceanliteracy.wp2.coexploration.org/next-generation-science-standards-2/>
- Offerdahl, E. G., Arneson, J. B., & Byrne, N. (2017). Lighten the Load: Scaffolding Visual Literacy in Biochemistry and Molecular Biology. *CBE- Life Sciences Education*, 16, 1–11. <https://doi.org/10.1187/cbe.16-06-0193>
- Paivio, A. (1991). Dual Coding Theory: Retrospect and Current Status. *Canadian Journal of Psychology*, 45(3), 255–287. Retrieved from <https://content.ebscohost.com/ContentServer.asp?T=P&P=AN&K=1992-07881-001&S=L&D=pdh&EbscoContent=dGJyMNHr7ESeqLI4y9fwOLCmr1Gep7JSsae4SrCWxWXS&ContentCustomer=dGJyMPGssU2uqrVMuePfgeyx33z01%2Bx74wAA>
- Quillin, K., & Thomas, S. (2015). Drawing-to-Learn: A Framework for Using Drawings to Promote Model-Based Reasoning in Biology. *CBE—Life Sciences Education*, 14, 1–16. <https://doi.org/10.1187/cbe.14-08-0128>
- Reece, J. B., Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., & Jackson, R. B. (2011). *Cambell Biology, 9th Edition*. San Franscisco: Pearson Benjamin Cummings.
- Saldaña, J. (2016). *The coding manual for qualitative researchers* (3rd ed.). Thousand Oaks: Sage Publications Inc.
- Santagata, S. (2011). Evaluating neurophylogenetic patterns in the larval nervous systems of brachiopods and their evolutionary significance to other bilaterian phyla. *Journal of Morphology*, 272(10), 1153–1169. <https://doi.org/10.1002/jmor.10975>
- Scheiter, K., Schleinschok, K., & Ainsworth, S. (2017). Why Sketching May Aid Learning From Science Texts: Contrasting Sketching With Written Explanations. *Topics in Cognitive Science*, 9, 866–882. <https://doi.org/10.1111/tops.12261>
- Schmidt-Nielsen, K. (1997). *Animal Physiology: Adaptation and Environment* (5th ed.). New York: Cambridge University Press.
- Schönborn, K. J., & Anderson, T. R. (2008). Bridging the educational research-teaching practice gap - Conceptual understanding, part 2: Assessing and developing student knowledge. *Biochemistry and Molecular Biology Education*, 36(5), 372–379. <https://doi.org/10.1002/bmb.20230>
- Schunk, D. H. (2012). *Learning Theories: An Educational Perspective* (Sixth). Boston: Pearson Education, Inc. Retrieved from https://bb.courses.maine.edu/bbcswebdav/pid-4718776-dt-content-rid-11699349_2/courses/1920.UMS05-S.43767./1920.UMS05-S.43767._ImportedContent_20190128020040/%5BDale_H._Schunk%5D_Learning_Theories_An_Educational..pdf

- Sezen-Barrie, A., Stapleton, M., & Miller-Rushing, A. (2020). Assessment of Students' Explanatory Models for Conceptual and Epistemic Quality: The Case of Ocean Acidification (OA) and its Impacts on Oysters. In *NARST 93rd Annual International Conference* (p. 294). Portland, OR: NARST.
- Strang, C., Decharon, A., & Schoedinger, S. (2007). Can You be Science Literature Without Being Ocean Literate? *The Journal of Marine Education*, 23(1), 7–9. Retrieved from www.coexploration.org/oceanliteracy/
- Townsend, D. W. (2012). *Oceanography and marine biology: an introduction to marine science*. Sunderland: Sinauer Associates, Inc.
- United Nations. (2018). *The Sustainable Development Goals Report*. United Nations Publications. <https://doi.org/10.18356/3405d09f-en>
- United Nations. (2019). World Oceans Day. Retrieved from <https://unworldoceansday.org/index.php/about>
- Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., & Reece, J. B. (2015). *Campbell Biology in Focus* (2nd ed.). Essex: Pearson Education Limited.
- Weatherbee, R. A., & Lindsay, S. M. (2018). Designing a Curriculum-Aligned Assessment of Cumulative Learning about Marine Primary Production to Improve an Undergraduate Marine Sciences Program. *Journal of Microbiology & Biology Education*, 19(3), 1–11. <https://doi.org/10.1128/jmbe.v19i3.1396>
- Webb, P. (2020). Nitrogen and Nutrients. Roger Williams University. Retrieved from <https://rwu.pressbooks.pub/webboceanography/chapter/5-6-nitrogen-and-nutrients/>
- Wiggins, G., & McTighe, J. (2005). *Understanding by Design* (2nd ed.). Alexandria: Association for Supervision and Curriculum Development.
- Wilkin, D., & Blanchette, J. (2019). Cnidarian Structure and Function. CK-12 Foundation. Retrieved from <https://www.ck12.org/biology/Cnidarian-Structure-and-Function/lesson/Cnidarian-Structure-and-Function-Advanced-BIO-ADV/>
- Wittrock, M. C. (1989). Generative Processes of Comprehension. *Educational Psychologist*, 24(4), 345–376. Retrieved from <https://content.ebscohost.com/ContentServer.asp?T=P&P=AN&K=6370283&S=R&D=a9h&EbscoContent=dGJyMNHr7ESeqLI4y9fwOLCmr1Gep7JSr6u4SbCWxWXS&ContentCustomer=dGJyMPGssU2uqrVMuePfgeyx33z01%2Bx74wAA>
- Wu, S. P. W., & Rau, M. A. (2019). How Students Learn Content in Science, Technology, Engineering, and Mathematics (STEM) Through Drawing Activities. *Educational Psychology Review*, 31, 87–120. <https://doi.org/10.1007/s10648-019-09467-3>

APPENDIX A

Table A.1. Codes developed and their descriptions. This table summarizes each parent code used as described in the Methods section, with an explicit description of the child codes contained within each parent code. Some child codes are scored on a 0 – 3 scale; others are more descriptive. Any sub-child codes (e.g. for the Alternative Conception code “Key Component Absent”) are also named and described under the “Sub-Child code and description” column. Codes appear in the table in the order in which they were used on notecard assessments.

PARENT CODE	PARENT CODE DESCRIPTION	CHILD CODE	CHILD CODE DESCRIPTION	SUB-CHILD CODE & DESCRIPTION
DEPTH OF DRAWING (DD)	This code gives a description of the level of complexity, as corresponding to scientific accuracy, interconnectedness of ideas, and thinking across scales of the drawing. This will indicate the level of detail and accuracy in a drawing that in turn reveals student understanding of the close physical integration of the circulatory and respiratory systems.	0	Drawing efforts do not capture even circuit-like nature of lungs, heart, and body interactions; may only have drawn piece-wise answer or no drawing whatsoever	
		1	“Baseline” drawing of a circuit diagram, with loops between lungs, heart, and/or body	
		2	Close investment of the two systems is represented by either a circuit diagram or diagram of lungs/gills alone that clearly depicts/labels capillaries.	
		3	Depiction of countercurrent exchange	
		N/A	Students were not prompted to draw a response.	

Table A.1. continued

PARENT CODE	PARENT CODE DESCRIPTION	CHILD CODE	CHILD CODE DESCRIPTION	SUB-CHILD CODE & DESCRIPTION
DEGREE OF VISUALIZATION (DoV)	This is a judgment of the notecard as a whole to determine to what degree students employ visualizations and drawings in their responses versus written text.	0	No visualizations used in response; purely written response to prompt.	
		1	Mostly words with minimal drawings; drawings that are included do little to support the ideas being described and/or depicted	
		2	Balanced combination of words and drawings.	
		3	More drawings and representations are employed than written text.	
KEY CONCEPTS (KC)	These are the key ideas that students should include in their notecard responses based on the information from the preceding lecture(s). Based on proximity of elements on a response for drawings (2019), items that are close to one another conveying a singular idea will be lumped and coded together, whereas items physically separated from one another will be coded singly.	MASS TRANSPORT (MT)	Drawing or explanation that clearly indicates that circulatory systems serve as mass transport systems that link gas exchange surfaces to cells in the tissue. The circulatory system moves O ₂ over distances that are too great for diffusion alone. In order for a circuit diagram to be considered displaying this concept, it MUST show a circuit that include the body and/or other tissues (not just heart and lungs/gills alone).	

Table A.1. continued

PARENT CODE	PARENT CODE DESCRIPTION	CHILD CODE	CHILD CODE DESCRIPTION	SUB-CHILD CODE & DESCRIPTION
KEY CONCEPTS (cont.)		CIRCULATORY AND RESPIRATORY SYSTEMS ARE PHYSICALLY INTEGRATED (CRI)	The integration of capillaries (circulatory system) into lungs or gills (respiratory system) allows for increased surface area for efficient gas exchange. Diagrams with clear depictions of capillaries into either of the respiratory structures, or written responses that clearly describe the role of capillaries in oxygen uptake, are considered to have demonstrated an understanding of this Key Concept.	
		COUNTERCURRENT EXCHANGE (CCE; BONUS)	(GILL examples only) Water flowing over gills during gill irrigation is countercurrent in direction as blood flowing in the capillaries of the gill lamellae; this allows for oxygen diffusion into blood over weaker concentration gradients than concurrent exchange for more efficient O ₂ extraction from the water.	

Table A.1. continued

PARENT CODE	PARENT CODE DESCRIPTION	CHILD CODE	CHILD CODE DESCRIPTION	SUB-CHILD CODE & DESCRIPTION
EVIDENCE	Co-coded with “Key Concepts” to demonstrate how students choose to display their knowledge in respond to the prompt, either in written or drawn format.	BOTH WRITTEN AND DRAWN	Evidence of Key Concept being conveyed is a combination of writing and drawing. Labels alone are not considered text/written responses; word descriptions must be in the form of captions or explanatory/descriptive statements to count as “written.”	
		DRAWING ALONE	Evidence is only drawn/ student only uses images to convey ideas. If the only text used to convey the Key Concept on the response is in the form of labels, the response will be coded as drawing alone if labels are simply pointing out features of a drawn representation without providing deeper explanation of the question prompt.	
		FLOWCHART	Student shows understanding of a Key Concept through arrows linking words alone or boxed words; this is not considered a drawn response.	

Table A.1. continued

PARENT CODE	PARENT CODE DESCRIPTION	CHILD CODE	CHILD CODE DESCRIPTION	SUB-CHILD CODE & DESCRIPTION
EVIDENCE (cont.)		WRITING ALONE	No attempt at drawing is made to convey knowledge about the Key Concept.	
		WRITING ALONE WHEN ASKED TO DRAW (2019 ONLY)	Students chose to communicate their knowledge about the Key Concept through writing alone; no attempt at drawing was made despite the explicit prompt to draw.	
ALTERNATIVE CONCEPTIONS (AC)	These are oppositions or shortcomings to a more scientifically accurate representation or explanation of the response. An Alternative Conception code does not necessarily mean the student is incorrect, but could instead indicate an inadequate response that does not explicitly demonstrate deeper thinking about the prompt.	COMPARISON	Circulatory and respiratory systems are compared to one another (physical similarities or similar processes depicted/described) rather than explaining how the two are linked.	
		UNCLEAR INTENT	Part or whole of response is unclear about what information or knowledge the student is trying to communicate.	
		“VITALISTIC” RESPONSE	Response explains linkage merely in terms of “vital signs” like breathing or heart pumping (e.g. “You need to breathe oxygen to keep your heart pumping”).	

Table A.1. continued

PARENT CODE	PARENT CODE DESCRIPTION	CHILD CODE	CHILD CODE DESCRIPTION	SUB-CHILD CODE & DESCRIPTION
ALTERNATIVE CONCEPTIONS (cont.)		RESPIRATORY INACCURACY	Some scientifically inaccurate representation or description of the respiratory system features and/or function(s) is present. For example, oxygenation may be inaccurately described (e.g. heart receives oxygen directly), or respiratory system parts may be left out of drawing or description entirely despite O ₂ being present.	
		CIRCULATORY INACCURACY	Some scientifically inaccurate representation or description of the circulatory system features and/or function(s) is present. For example, capillaries may be labelled as veins or arteries, or a student may describe gill capillaries as carrying water instead of blood.	

Table A.1. continued

PARENT CODE	PARENT CODE DESCRIPTION	CHILD CODE	CHILD CODE DESCRIPTION	SUB-CHILD CODE & DESCRIPTION
ALTERNATIVE CONCEPTIONS (cont.)		TERMINOLOGY	Confuses terminology in written descriptions by misusing words (e.g. “water is directly respired”) or vague explanations (e.g. “maintaining a level via diffusion”). Student may also use a term without adequate explanation (e.g. just writes “countercurrent exchange” without any deeper description or depiction).	
		KEY COMPONENT ABSENT	Drawing or explanation lacks critical component of system for explanation/ response to be considered complete. Alternatively, a microscopic process (such as countercurrent exchange, CCE) is depicted in isolation without a linkage to macroscale structures or processes.	Four sub-child codes exist to specify <i>which</i> key component is absent: (1) CCE depicted in isolation; (2) capillaries absent; (3) only circulatory system described; and (4) only respiratory system described.
REPRESENTATIONS	This is a list of different recurring elements seen in drawn responses, or explicitly described or named in written responses.	PROCESSES	A single sub-child code exists: gas exchange/diffusion.	Expressly draws, labels, names, or describes the process of gas exchange occurring between the two systems.

Table A.1. continued

PARENT CODE	PARENT CODE DESCRIPTION	CHILD CODE	CHILD CODE DESCRIPTION	SUB-CHILD CODE & DESCRIPTION
REPRESENTATIONS (cont.)		CLARITY OF LINKAGE	This is an assessment of linkers used to describe the relationship between the two systems.	Three sub-child codes exist: (1) arrows/words describing directionality to indicate blood and/or air/water flow; (2) arrows/words indicating a relationship between structures described or depicted; and (3) arrows/linkers absent, in which case separate explanations or drawings are present with no clear linkages between the two systems.
		STRUCTURES	These are physical structures or objects that students describe or draw in their responses.	Six sub-child codes exist in this category: (1) heart, (2) lungs/gills, (3) capillaries, (4) pulmonary circuit (heart and lungs/gills only), (5) pulmonary + body circuit, and (6) written molecule (CO ₂ and/or O ₂)
		SYSTEM-LEVEL DEPICTION	"Circulatory system" or "respiratory system" discussed or depicted only, with no finer level of detail	

Table A.1. continued

PARENT CODE	PARENT CODE DESCRIPTION	CHILD CODE	CHILD CODE DESCRIPTION	SUB-CHILD CODE & DESCRIPTION
REPRESENTATIONS (cont.)		DRAWINGS ONLY	This category of representations only applies to drawn responses.	Two sub-child codes exist: (1) Slide diagram, in which a student's drawing is a clear attempt at reproducing a diagram that was shown in the lecture slides; and (2) Unessential drawing, in which the entire drawing/aspect of the drawing or diagram is does nothing to support the student's answer.
COHERENCE (COH)	This is an assessment on how seamless a student explanation is in response to the prompt, scored after KC, AC, and Representations are evaluated. This score looks for grammatical and/or diagrammatic linkers between ideas in order to assess if there are any "gaps" between student thinking. This will be noted in the upper right-hand corner when possible and is separate from the DoV and AD boxes.	0	Response has no connections/relationships among ideas. Ideas may be present just in list form but no linkers to show how they are related. Major Alternative Conception(s) present such that the response does not address the prompt.	

Table A.1. continued

PARENT CODE	PARENT CODE DESCRIPTION	CHILD CODE	CHILD CODE DESCRIPTION	SUB-CHILD CODE & DESCRIPTION
COHERENCE (cont.)		1	Response has minimal connections or relationships between ideas. Contains partly accurate information, but may communicate chain of ideas with only partial linkage between them. May have one or more ACs present such that they cloud accuracy of response.	
		2	Response has moderate connections or relationships between ideas. One or more KCs present, but may lack linkers between ideas to provide “gapless” thinking/reasoning. Any ACs that are present are minor, but inclusion of these indicates a dissonance in expertise of knowledge.	
		3	Response demonstrates extensive connections and relationships among ideas; seamless explanation/depiction of ideas with no gaps in thinking. The sequencing of ideas makes logical sense and no ACs present.	

Table A.1. continued

PARENT CODE	PARENT CODE DESCRIPTION	CHILD CODE	CHILD CODE DESCRIPTION	SUB-CHILD CODE & DESCRIPTION
EXPERTISE SCORE (ES)	This gives a score from 0-3 that assesses the degree to which the student response aligns with an “expert” answer. This will indicate the overall level of accuracy, thoroughness, completeness, and clarity of the response as a whole, and will be a mechanism that will allow us to compare the quality of ALL responses, be they written or drawn. This score goes beyond Coherence to also assess expert thinking qualities that include: interconnected ideas, thinking across biological scales, and systems-based thinking. The ES code will be placed in the bottom righthand corner of the card if space permits; otherwise it will appear in the bottom left.	0	Prompt not addressed; only circulatory or respiratory system is described or depicted at some level, but not both. A comparison between the two systems may have been made rather than addressing the linkage(s) between the two.	
		1	Novice; both systems are described or depicted in some way, but there is explicit understanding of linkage described that addresses the prompt specifically.	
		2	Medium; connection between respiratory and circulatory system shown to some degree, but may not be described or depicted <i>explicitly</i> with capillaries, physical proximity/ integration, and/or that blood is the means of transporting oxygen over long distances that cannot be covered by diffusion alone. Countercurrent	

Table A.1. continued

PARENT CODE	PARENT CODE DESCRIPTION	CHILD CODE	CHILD CODE DESCRIPTION	SUB-CHILD CODE & DESCRIPTION
EXPERT SCORE (cont.)		3	Expert; linkage between systems clearly and thoroughly described/depicted AND contextualization of countercurrent exchange in fish examples (if lungs depicted, not necessary to include).	

exchange may be depicted but not contextualized.

BIOGRAPHY OF THE AUTHOR

Christina Siddons (née Dykeman) was born in Hudson, New York in 1988 and graduated from Albany Academy for Girls in Albany, New York in 2006. She went on to attend Cornell University in Ithaca, New York, where she fell in love with sailing and marine fieldwork during her semester abroad with Sea Education Association (SEA) in the fall of 2008. With SEA, she sailed for six weeks aboard the tall ship *SSV Robert C. Seamans* from San Diego, California to Puerto Vallarta, Mexico, where she studied pteropods, a type of pelagic zooplankton. She graduated from Cornell in May of 2010 with a Bachelor of Science in Biology degree from the College of Agriculture and Life Sciences.

Following her graduation from Cornell, Chrissy followed whatever path took her to the ocean, which resulted in a research apprenticeship in the San Juan Islands, a season working as a marine educator at Salish Sea Expeditions in Puget Sound, and finally a return to SEA, this time for a five-year stint as an Assistant Scientist. During her time at SEA, Chrissy sailed 45,000+ nautical miles all over the world, including the waters off Cape Cod, the Caribbean, and the South Pacific. Chrissy moved to Maine in the fall of 2015; she worked at the MDI Biological Laboratory as a research assistant for three years, investigating the genetic regulation of heart regeneration in zebrafish. Although she enjoyed the research, she missed teaching, which was her main focus at SEA. Chrissy therefore decided to pursue obtaining her teaching certification to teach middle school math and science, and hopefully bring her joy of the oceans to classrooms around Downeast Maine. She is a candidate for the Master of Science degree in Teaching from the University of Maine in August 2020.