2020

Semester-Long Instruction In Drawing For Biology Changes Study Habits, Motivation To Draw, And Approaches To Problem-Solving

Jessica Rose Burns
William & Mary - Arts & Sciences, jessicaroseburns@hotmail.com

Semester-long Instruction in Drawing for Biology Changes Study Habits, Motivation to Draw, and Approaches to Problem-solving

Jessica Rose Burns

Meisertown, Pennsylvania

Bachelor of Science, East Stroudsburg University, 2018

A thesis presented to the Graduate Faculty of The College of William & Mary
in Candidacy for the Degree of
Master of Science

Department of Biology

College of William & Mary
May 2020

APPROVAL PAGE

This thesis is submitted in partial fulfillment of
the requirements for the degree of

Master of Science

Jessica Rose Burns

Approved by the Committee April 2020

Committee Chair
Paul Heideman, Professor, Department of Biology
College of William & Mary

Drew LaMar, Professor, Department of Biology
College of William & Mary

Jason Chen, Associate Professor
William & Mary School of Education

COMPLIANCE PAGE

Research approved by

Protection of Human Subjects Committee

Protocol number(s): PHSC-2018-08-23-13077-pdheid

PHSC-2019-08-05-13782-pdheid

Date(s) of approval: 23 August 2018

20 September 2019

ABSTRACT

Model building can use drawing or sketching as a mechanism to help the drawer learn information (study), solve problems (model-based reasoning), and communicate. Unfortunately, many students fail to master drawing or sketching skills due to the effort and instruction required. Additionally, few longitudinal, real-world classroom studies have been conducted on the teaching of drawing to students. We applied guided practice of drawing or sketching to an undergraduate first-semester Introductory Biology majors course, aiming to assess (1) the patterns of growth and decline in the use of sketching and other active study methods over subsequent semesters, (2) the relationship between usage of sketching by students and performance, and (3) student motivations (self-efficacy, utility value, interest, and cost) and attitudes towards drawing and sketching as a learning tool. Students with instruction on drawing as a learning tool decreased their use of passive study methods during the course and increased their use of active methods. Major changes included less rereading in studying and more drawing or sketching. One semester after the course, these students maintained part of the gains in drawing and active study methods, using both significantly more than prior to the drawing intervention. Students without the instruction in drawing showed few changes during the two semesters. Higher proportions of study time spent drawing predicted higher overall course point total. Students with instruction on drawing reported higher self-efficacy towards drawing. However, only cost value predicted use of drawing during study time, suggesting that instructors interested in teaching drawing as a learning tool should aim to decrease perceived cost for the students. These outcomes will be reassessed yearly. Our preliminary conclusion is that a course in this format can support development of drawing or sketching for learning while developing more active study methods.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to thank my research advisor, Dr. Paul Heideman for giving me the chance to take on such an enormous research project and for supporting me as I learned new techniques and material. My committee, Dr. Drew LaMar and Dr. Jason Chen were also irreplaceable resources for multidisciplinary work, guiding me to resources and reviewing my work. To the students who dedicated themselves to learning new topics and diligently worked on my research team (Emma Arents, Nicole Harris, Niko Lignore, Catie Lott, Alex Tan, Emma Watson, and Maddy Zamecnik): this project would not be the same without your curiosity and I hope we can learn together as this project continues in the years to come. The final thank you goes to my mother, Karen Trebitz, and my husband, Felipe Sottorff Araya, for letting me bounce ideas off them and for patiently listening to my rants about research… "I don't believe in humans!"

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Semester-long Instruction in Drawing for Biology Changes Study Habits, Motivation to Draw, and Approaches to Problem-solving

1. Introduction

1.1 Study Skills, Drawing, and Guided Practice

Past studies exploring the effectiveness of different study methods demonstrate that those most commonly used by students, such as rereading and highlighting, are among the least effective (Dunlosky et al., 2013; Fu and Gray, 2004). Additionally, there is research that indicates that students often fail to adopt alternative problem-solving techniques and study methods even if they know alternatives are more effective than their current method (Pressley et al., 1989; Donovan and Bransford, 2005; Fu and Gray, 2004). Students also struggle to organize data in a way that supports reasoning using mental or visual models (Quillin and Thomas, 2015; Dauer and Long, 2015).

Models are mental or physical representations of an object, process, or idea. Models can be two or three dimensional and they may be static or include movement. Models are typically smaller and simpler than real structures or systems. For example, a stick figure drawn by a child clearly indicates a human, while leaving out the majority of the details. This simplification allows large and complex sets of information to be represented and accessed efficiently. Additionally, models can be used to solve problems through simplifying complex ideas into a manageable format, a skill referred to as model-based reasoning (Quillin and Thomas, 2015; Dauer and Long, 2015). Model-based reasoning is used to work through problems in a variety of fields, including mathematical formulas, chemical reactions, or a diagram of a cell in biology.

Modelling is important because science, technology, engineering, and mathematics (STEM) disciplines often require the ability to use visual representations to communicate ideas, as well as a deep enough understanding to convert between visual and verbal explanations of processes (Gilbert, 2005; Mathewson, 1999; cited in Wu and Rau, 2018; cited in Quillin and Thomas, 2015). Unfortunately, this is a skill that students struggle to obtain early enough in their careers.

Van Meter and Garner's (2005) literature review on learner-generated models summarized evidence that drawing is an effective way for students to interconvert between internal models and external communication. They found that drawing could be used to integrate new information into internal models, organize topics within these internal models, and select pertinent information from internal models. Therefore, drawing could be used to internalize new information and to express knowledge contained within the drawing, as well as to perform model-based reasoning (Quillin and Thomas, 2015). However, as with many learning and communication tools, effective drawing skills can be difficult and time-consuming for students to learn on their own.

There is research suggesting that students can learn necessary skills if provided with the appropriate learning environment and toolkit. Students applying more active study methods and learning to self-monitor the effectiveness of their learning habits (metacognition) can improve overall learning (Bielaczyc, Pirolli and Brown, 1995). Additionally, research summarized by Van Meter and Garner (2005) suggests that students will learn drawing skills faster and better when they are allowed guided practice. This guided practice involves the presentation of a task or skill by the instructor (the expert) and opportunities for the students (the beginners) to practice this skill, followed by consistent instructor feedback on the quality of students' work involving the skill they are trying to master.

A previous guided practice study by Heideman et al. (2017) indicated that using 'minute sketches' (which are simple, quickly made drawings of one's mental model, including only enough detail to make sense to the person drawing it) improved retention of novel biology topics and problem-solving on exams. About eight months after learning 'minute sketching', participants reported using about 25% of their study time drawing and redrawing, compared to about 8% of study time in the comparison group; this indicates that a short intervention to teach sketching might lead to higher adoption rates by students. However, when re-surveyed two years later, the difference in sketching usage between the intervention 'minute sketching' group and the comparison group had evaporated. In contrast, students who had taken one or more courses that included (1) extensive practice in drawing, and (2) used sketches for model-based reasoning, often in course assignments, applied sketching more often in their later courses than those who had not, suggesting that longer exposure with practice may help retain the use of drawing-to-learn in the long-term (Heideman et al., in preparation). Unfortunately, few studies have applied methods to develop drawing skills and assessed long-term application by students of drawing for learning and model-based reasoning in a real-life classroom (reviewed by Quillin and Thomas, 2015). We have found none with continued tracking over months to years for applications of sketching for long-term recall of systems and concepts with model-based reasoning.

1.2 Theoretical Framework

Studying plays an essential role in a college education. Past research exploring the effectiveness of different study methods demonstrates that those commonly used by students, such as rereading and highlighting, are among the least effective (Dunlosky et al., 2013; Fu and Gray, 2004). It might be that these students just do not know better, that they have not been taught more effective methods, or they do not know how effective these methods can be. However, even when students have been taught which methods are more effective, students typically still use the less effective study methods. So why do students continue to use study methods they know to be less effective?

In order to approach this question, we must first understand the theories behind how students learn and process information. A theory applicable for sketching regarding learning is Richard Mayer's cognitive theory of multimedia learning (CTML). CTML, simply put, states that by matching learning material to how the human brain processes information, one can increase meaningful learning.

According to CTML, the human brain processes visual and auditory stimuli differently (Mayer, 2003, 2005, 2009). More specifically, for learning to occur, these types of stimuli must travel through different processing paths. Visual information, such as a picture, is observed through the eyes (sensory memory), and from it select information is organized into a mental model (working memory), before being integrated with prior knowledge (long-term memory). Auditory information, such as spoken words, starts by entering the ear (sensory memory) before following an identical, but parallel, path to visual stimuli. This parallel processing leads to three assumptions: **dual channels, limited capacity, and active processing.**

Dual channels, the first assumption, is that information from a stimulus may pass through one of two paths through the brain, one visual/pictorial, and the other auditory/verbal (Mayer, 2003, 2005, 2009). Either or both must carry information to be committed to memory. This parallels Paivio's dual-coding theory (1990, 2007; Clark & Paivio, 1991) and Baddeley's working memory model (1992) for two separate and distinct pathways to transfer information to memory. The second assumption, **limited capacity**, is that working memory can only contain a few items at a time, therefore putting strict limits on the amount of information one can process at a given time. This parallels Sweller's cognitive load theory, which suggests that limiting the overall amount of information received at one time increases one's ability to process and retain information (1999, 2005; Chandler and Sweller, 1991). Past research shows that people can hold 5-7 items or chunks in their working memory at one time (Miller, 1956; Simon, 1974). The final assumption, **active processing**, implies that the learner is part of, and aware of, the processing of information as it occurs (Mayer, 2005, 2008, 2009). Specifically, the learner can choose what to pay attention to, which is referred to as “selecting”. Learners “organize” this selected information in their minds before “integrating” the information by tying it into their previous knowledge.

The first assumption governing CTML implies that two flows of information can be processed at one time if one stimulus is auditory and the other is visual. Combining these paths allows for faster processing and increased learning. But what happens when you have limited capacity, the second assumption, and multiple stimuli competing for your attention?

To address this, Mayer assessed factors that may affect the effectiveness of multimedia learning. Based on a combination of his own empirical studies and those he reviewed, Mayer outlined twelve principles of multimedia learning, which define factors to consider when presenting learning material (Mayer, 2009). Although all of these are important to learning gains, we focus on the four principles most relevant to the drawing-to-learn methodology used in our study: **coherence, redundancy, spatial contiguity, and pre-training**.

The **coherence principle** states that decreasing extraneous information increases learning (Mayer, 2003, 2009; Figure 1A). The simple sketches taught in our method are by definition minimalistic; they are formed from simple shapes, and non-essential details are left out to allow the learner to focus on important details. The **redundancy principle** states that removing repeated information increases learning (Mayer, 2009; Figure 1B). The sketches in our method require simple labels, with no full description, and students are expected to leave out repeated information such as structural subunits of a cell membrane -- they may instead just draw an arrow or line that indicates the pattern continues. The **spatial contiguity principle** states that related information should be placed in close proximity (Mayer, 2003, 2009; Figure 1C). Our sketching method requires students to label important parts in a clear spatial relationship to sketches in order to help associate words with sketches (Figure 1E). Lastly, the **pre-training principle** says that providing students with basic building blocks prior to a lesson decreases cognitive load, thereby increasing their learning (Mayer, 2009; Figure 1D). Students taught our method of sketching were provided with instruction on how to create sketches, given examples, and informed on the types of items to include before constructing their own sketches. We applied this framework to encourage students to develop sketches that are information-rich but can be drawn, labeled, and practiced quickly.

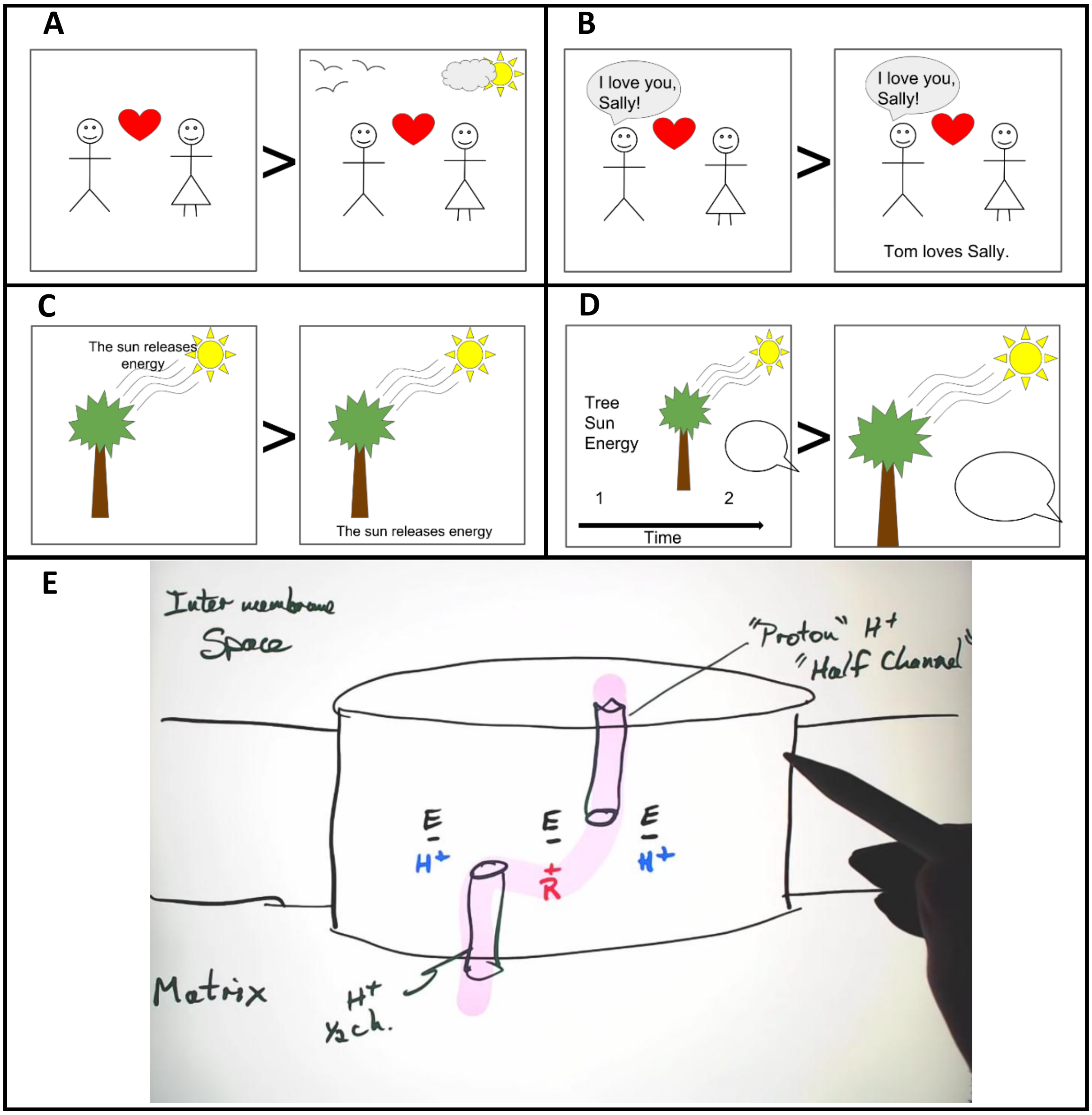


Figure 1 – **Illustrations based on Mayer's principles of multimedia learning.** Demonstrated are versions of the same material that follows a given principle (left side) compared to one that fails to follow each principle (right side): (A) coherence principle, (B) redundancy principle, (C) spatial contiguity principle, and (D) pre-training principle. (E) shows an example sketch from a YouTube video (Paul Heideman 2018) posted by the professor as part of the homework assignments provided to the students (see Methods) that demonstrates the first 3 of these principles.

1.3 Motivation

An understanding of how students learn and process information, as well as the things that help assist or act as barriers to these processes, allows instructors to teach students effective methods to represent structures, events, and concepts in sketches. Students can be shown why these methods are effective, and they can learn to apply an instructor's sketch or their own in order to solve a problem. Why, then, do students fail to adopt more effective study methods even when they know better? The problem may be one of motivation: the reason a person acts in a certain way.

Expectancy-value theory splits motivation into two factors, expectancy and value, to help us understand some of the factors influencing actions such as students' usage of our drawing-to-learn study method (Wigfield and Eccles, 2000; Fishbein, 1975). Fishbein (1975) and Wigfield and Eccles (2000), concluded that people's actions are caused by the combination of the degree to which people expect to succeed at a task and how much they value the activity and its outcome. These factors affect performance, achievement choices, effort, and persistence (Wigfield and Eccles, 2000). Since these factors can influence students' actions, it is important to quantify student motivation towards the topics and tools used in the classroom. Our interest is in how students view sketching as a learning tool.

In our study, we chose to look at one measure of expectancy (self-efficacy), and three measures of value (cost, interest, and utility). Self-efficacy reflects how confident students are in their ability to achieve a specific goal, such as sketching a quick and simple drawing to illustrate a scientific process (Bandura, 1997). Self-efficacy is important in academics because students are more likely to avoid a task if they do not believe they can do it (Schunk and DiBenedetto, 2016). However, self-efficacy can be increased by modeling, having another person demonstrate, the skill (Schunk, 1989; Schunk, 2012). Higher self-efficacy leads to an increase in effort, metacognition, competence, and persistence (Lent et al., 1984; Lent et al., 1986; Pintrich and Garcia, 1991; Pajares, 1996, Usher and Pajares, 2008). Increased persistence can lead to an increase in performance on memory tests (Berry, 1987) and an increase in academic performance (Lent et al 1984; Lent et al 1986). A meta-analysis by Multon et al. (1991) found that self-efficacy accounts for about 14% of the variation in academic performance (Schunk and DiBenedetto, 2016). Self-efficacy also acts as a mediator between past and future performance, therefore acting as a stronger predictor of future performance than the prior performance itself (Usher and Pajares, 2008). Higher self-efficacy can improve problem-solving skills (Bouffard-Bouchard, 1990; Usher and Pajares, 2008).

Interest value is related to students' feelings towards an experience or skill. Interest comes in two main forms: situational interest and individual interest (Hidi and Renninger, 2006). Situational interest is temporary interest, often caused by something, or someone, in the environment. This could be a catchy video shown by a teacher. Individual interest, on the other hand, comes from a person's direct interest in the topic. Situational interest can become individual interest over time, should the person continue to be interested in the topic after the original cause passes. Generally, students learn better if they consider the topic or material to be more interesting (Pintrich and Schunk, 2002). However, the type or level of learning may vary; Mayer et al. (2008; Pintrich and Garcia ,1991; Pintrich and Schunk, 2002) found that an increase in interestingness often led to an increase in retention of provided information, but also led to a decrease in transfer scores. This suggests that higher interest may capture and hold students' attention on a topic, but it does not help students apply what they learn to novel situations. Interest value correlates more with learning gains in low achieving students (Hulleman and Harackiewicz, 2009), suggesting that attention grabbers and other ways to increase interest can be used to help these students improve their academic performance.

Utility value and cost value consider the intrinsic, or internal, motivation of students in relation to external factors in their lives. These could be considered pros and cons, or benefits and costs, to the students. Utility value is whether students find an experience or skill relevant or useful (Wigfield and Eccles, 2000). For example, a student may find one technique helps them learn material better than another, and therefore they choose to use it more. Higher utility value predicts an increase in both interest value and academic performance (Hulleman et al., 2010). In contrast, cost value reflects how much students feel an experience or skill is worth. Cost can be viewed from many angles, including money, time, and effort (Wigfield and Eccles, 2000). If performing a task is too costly, it decreases the likelihood of it occurring. Higher utility and lower cost can both lead to increased performance (Rosenzweig et al. 2020).

1.4 Research aims

Even though drawing as a learning and problem-solving tool has been found to be effective, motivation may be central to explaining why the majority of undergraduate students fail to adopt this tool. Expectancy-value theory helps us understand some of the factors influencing actions related to application of the drawing-to-learn study method. One's actions are caused by the combination of the degree to which a person **expects** to succeed at a task and how much they **value** the activity and its outcome (Wigfield and Eccles, 2000; Fishbein, 1975). These factors affect performance, achievement choices, effort, and persistence (Wigfield and Eccles, 2000). It is therefore important to quantify student motivation towards the use of sketching as a learning tool.

We applied guided practice of drawing to an undergraduate first-semester Introductory Biology majors course, aiming to assess: (1) the patterns of growth and decline in the use of sketching and other active study methods over subsequent semesters, (2) the relationship between usage of sketching by students and performance (per question exam scores, remembering facts about a topic, problem solving related to Introductory Biology topics), and (3) student motivations (self-efficacy, utility value, interest, and cost) and attitudes towards drawing and sketching as a learning tool. This thesis presents initial results from the first 20 months of a 5-year longitudinal study on the use of drawing and student motivation in using drawing as a method for learning as students continue their education.

2. Methods

2.1 Course Design

2.1.1 Cohort with drawing intervention

Interview participants were students enrolled in 2 sections of BIOL 203, Introductory Biology: Cells and Molecules and Development, during the Fall 2018 semester at the College of William and Mary. In this course, students were introduced to sketching in a format intended to follow the guidelines suggested by the background literature (Van Meter and Garner, 2005, etc.), including (1) instruction and demonstration of drawing using the course material (2) demonstration of model-based reasoning using drawing, including talking through the problem solving process, (3) requirement of student practice through assigned exercises and homework problems, (4) followed by instructor feedback. This format was intended to help them (a) practice, (b) remember, (c) understand, and (d) problem solve. Students received instructional input on how to sketch and opportunities for practice with occasional feedback from the professor. They watched the instructor sketch class material during lectures and watched videos of sketches being drawn with explanatory voiceover (i.e. Paul Heideman 2018). These sketches represented structures, sequences of events, and concepts, all presented in a minimalist style, usually as “minute sketches” (Heideman et al., 2017). For homework, students produced guided sketches based on the assigned videos, or generated their own sketches of a topic discussed in class. As part of the introduction to sketching for learning, the professor also provided information on effective learning and studying, and explained applications of sketching for model-based reasoning.

2.1.2 Cohort without drawing intervention

Interview participants were students enrolled in 2 sections of BIOL 203, Introductory Biology: Cell and Molecular, during the Fall 2019 semester at the College of William and Mary. In this course, students received the same biology material as the 2018 class, except they had different professors and were not taught sketching nor heavily encouraged to draw as part of their coursework.

2.2 Data Collection

This research was approved and governed by a human subjects protocol entitled “Changes in Study Methods after Courses Including Instruction in Sketching” (PHSC-2018-08-23-13077-pdheid, PHSC-2019-08-05-13782-pdheid).

Data was collected using a combination of surveys sent via email to participants and in-person interviews. Surveys with an 8-12 minute typical response time addressed the aims related to study methods and use of sketching. This series of surveys tracked students in order to follow their study habits, including sketching, over their undergraduate career. The interviews aimed to answer the remaining research questions, focusing on the thought process of students while problem solving on exams. Each survey and interview is discussed in more detail in the following sections.

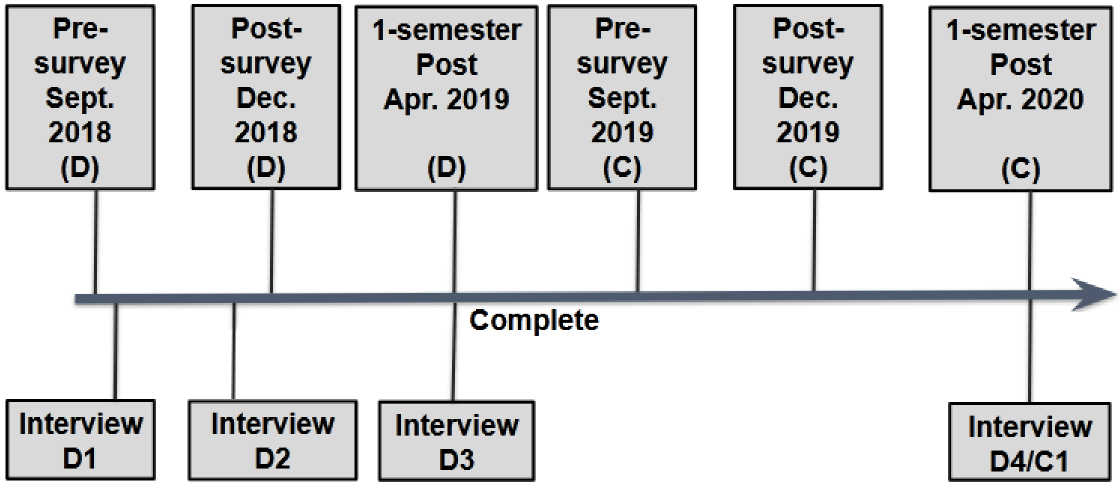


Figure 2 - **Timeline of surveys and interviews**. Only completed timepoints are shown (Fall 2018-Spring 2020), as these are the data discussed in this paper. “D” represents timepoints for the cohort with the drawing intervention. “C” represents timepoints for the cohort without the drawing intervention. Only the cohort with the drawing intervention participated in interviews during the Introductory Biology course.

2.2.1 Pre-survey

In September of the first year of each cohort (drawing and without drawing intervention), the pre-survey was pushed to all students (n_{2018} = 219; n_{2019} = 458), regardless of academic social class (Freshman, Sophomore, etc.). For the cohort with the drawing intervention (2018) this included students who were currently enrolled in the two sections of BIOL 203 that were taught by instructor PH. For the cohort without the drawing intervention (2019) this included all students then enrolled in the two sections of BIOL 203, both co-taught by LA and SH. The Qualtrics™ link was distributed by the professor teaching BIOL 203 in 2018 (students were informed that the instructor would see no responses nor be informed about which students participated in the research until after grades were assigned at the end of the semester) and by the primary student researcher in 2019. For the Fall 2018 cohort (with drawing intervention), the survey was pushed three times and received a total of 180 responses out of the 219 enrolled students. For the Fall 2019 cohort (without drawing intervention), the survey was pushed three times and received a total of 128 responses of the 458 enrolled students. These responses were sorted using the optional study code provided by participants to remove obvious duplicates. Responses were checked to ensure they were answered completely and according to the directions. Since responses were optional, there were also questions without responses. Incomplete or missing responses did not disqualify the survey respondent as a whole, but instead removed that respondent from any analysis about a given question.

Due to additions to our research questions, the Fall 2019 cohort (without drawing intervention) received an alternate survey than the Fall 2018 cohort (with drawing intervention). In addition to questions found in the original course surveys, these surveys include questions regarding student motivation to draw. These questions are discussed in more detail in the section 2.2.3 “Follow-up surveys”, as that timepoint was used for the analyses.

2.2.2 Post-survey

In December of each cohort (drawing intervention and without drawing intervention), a second survey was pushed by the professor (2018) or the primary student researcher (2019) to all students. For the Fall 2018 cohort (with drawing intervention), the survey was pushed 3 times and received a total of 107 responses. This survey included many of the same questions as the original, but also a few on the structure of the course. For the Fall 2019 cohort (with drawing intervention), the survey was pushed 3 times and received a total of 91 responses. Again, due to additions to our research questions, the Fall 2019 cohort (without drawing intervention) received a different survey from the Fall 2018 cohort (with drawing intervention), including additional questions regarding student motivation to draw. These underwent the same cleaning process as the pre-surveys.

2.2.3 Follow-up surveys

The April after their course (one-semester after the post-survey), the students from each cohort (drawing and without drawing intervention) received a follow-up survey. These surveys included the additional questions regarding student motivation to draw. The surveys were identical for each cohort except for an additional question for the Fall 2018 cohort (with drawing intervention) asking about previous participation in interviews (see supplementary materials for a copy of this survey).

The follow-up surveys for both cohorts included questions regarding student motivation to draw. More specifically, they asked questions related to the constructs of self-efficacy, utility value, interest, and cost. The self-efficacy construct consisted of five items with a Cronbach’s alpha of 0.810 for the cohort with the drawing intervention and 0.868 for the cohort without the intervention. Self-efficacy can be measured with questions such as “How confident are you that you can draw a complex process (such as translation) from your course material?”. The interest construct consisted of three items with a Cronbach’s alpha of 0.917 for the cohort with the drawing intervention and 0.897 for the cohort without the intervention. Interest can be assessed using statements like “I think drawing to learn is enjoyable.” The cost construct consisted of three items with a Cronbach’s alpha of 0.761 for the cohort with the drawing intervention and 0.713 for the cohort without the intervention. This can be assessed using statements like “Drawing to learn is not worth all the time and effort it takes.” The utility construct consisted of three items with a Cronbach’s alpha of 0.807 for the cohort with the drawing intervention and 0.817 for the cohort without the intervention. Utility value can be assessed using statements such as “Drawing to learn is useful to my professional goals.” Depending on the construct, students responded based on a 5- or 7-point Likert scale system.

2.2.4 Requesting Interview Participants

After the add/drop period for the Fall 2018 semester passed, we identified Freshmen from the final class roster based on the listed graduation year; this risked some participants having previous university level experience, but was the most accurate way to filter for freshmen without individually asking the status of all 219 BIOL 203 students. This subset was randomized in R using an arbitrarily generated seed. Based on this randomized list, students were invited via email to participate in the interview portion of the study. Answers were recorded as “yes”/“no”/“no response” for each. “Yes” students were thanked for their interest and told they would receive more information at a later date. “No” students were thanked for their consideration and their offer was extended to the next student on the list. “No response” students received a second copy of the email after 3-4 days, requesting that they respond, “even if your answer is no.” After a week, “no response” students were considered a “no”, and therefore their offer was extended to the next student on the list. The desired sample size, 60 students, was chosen such that typical numbers of students who choose to withdraw from the study would still leave sufficient data for analysis. However, we only reached a true sample size of 42 participants for the first round of interviews. Therefore, this sample size was used as the maximum for all remaining interviews of this subset of students.

2.2.5 Interview Format

Interviews followed a modified version of Emily Hauge’s (now Gericke) interview protocol (Hauge, 2018), which was based on “think aloud” interviews (Hmelo Silver, 2004; Dauer and Long, 2015) and semi-structured interviews (Dye and Stanton, 2017).

Participants were asked to walk the interviewer through their thought process as they solved several problems on the exam. This included up to two free response and two multiple choice questions per exam. In addition to these four questions, students were asked about one process or concept important to introductory biology. Not all questions were covered in all interviews because interviews ended after 25 minutes.

Interviews were conducted in the week following the return of graded exams to allow time for exams to be graded and returned to students. Since we did not interview students on the day of the exam, we tried to minimize the time gap to allow for more accurate recall of their thought process while taking the exam. We recorded the number of days between the exam and the interview.

2.2.6 Interviews during the Course with Drawing Intervention

The first round of interviews was conducted during the week after the first midterm exam, from October 1-7. An anonymized Doodle Poll was emailed to students who agreed to participate in the study, allowing them to choose a time for their interview; 42 students completed this interview. Depending on the time chosen, interviews were assigned to one of four interviewers. The assigned interviewer confirmed the interview a day before, as well as requesting participants to complete the Qualtrics™ pre-survey if they had not already done so.

Exam 1 interviews began with a brief pre-interview information session. This included signing of consent forms (see Appendix A1 “Consent Form Example”), a brief explanation of participation requirements and the interview process, and a request of participants to provide us with what they remember of their survey study code. This code allowed us to link interviewees to their survey responses.

The second round of interviews occurred after the third midterm exam, which was on November 19. Interviews were conducted from November 28 -

December 7. This longer period between the exam and the interviews was due to a break in the academic calendar. These interviews followed the same interview protocol as Exam 1 interviews, except the pre-interview information session contained signing of the “Financial Operations Human Subjects Payment Log (less than $50)”, confirmation of study codes, and a reminder to complete the post-survey in the following weeks. One student dropped out of the interview portion of the study before completing the second interview, therefore 41 students completed this interview.

2.2.7 Follow-up Interviews

Additional interviews were conducted with students from the Fall 2018 cohort (with drawing intervention) of the BIOL 203 course, as well as students from the Fall 2019 cohort (without drawing intervention). As with previous interviews, students were asked to describe a common biological process learned in BIOL 203, such as transcription and translation. These interviews differed in that students also discussed their thought processes as they solved problems from recent exams they chose to bring to the interview, rather than from designated exams. Each student was asked to choose an exam from their current courses in that spring semester that the student found challenging (not “easy” questions) and the student felt represents their skills at problem solving OR that the student felt could have been solved multiple ways. Allowing students to choose from their current coursework guaranteed recent material and allowed for topics other than biology to be discussed; in turn, this allowed for more generalizability. As with other interviews, the focus was on their process, not the correctness of a response.

In April 2019, 30 students from the original drawing cohort completed a one-semester post-course interview. These students discussed one or two exam questions of their choice from the exam they brought to the interview. Additionally, these students were asked to describe the process of transcription aloud to the interviewer. They were given paper allowing them to write notes or draw prior to or while responding, but participants were not required to use the paper provided. They then solved an application problem related to transcription factors on paper, for which they were allowed to respond in words, images, or a combination of the two. Once participants felt they had completed the question, they were asked to explain their answer and thought process in the same format used to discuss exams from their current courses.

In April 2020, students from the original cohort without the drawing intervention completed a one-semester post-course interview. These students completed the same one-semester follow-up interview process and questions as the cohort with the drawing intervention, except that interviews were conducted via Zoom due to the COVID-19 campus closure. However, these student interviews were not included in this thesis.

2.3 Data Analysis

2.3.1 Research Question 1: Patterns of growth and decline in the use of sketching and other active study methods

Quantitative analyses of surveys were completed to test the predictions that (a) students who have taken the class with a drawing intensive format will use drawing and other active methods as a greater portion of their study time than students who have not, and (b) students will use drawing and active methods as a greater portion of their study time at the end of the course than at the beginning. These analyses followed similar methods to those used by Heideman et al. (2017), with individual t-test comparisons between cohorts at each timepoint and between timepoints within each cohort. Due to the multiple statistical comparisons conducted on each dataset, we used the false discovery rate control to adjust significance (Benjamini and Hochberg, 1995; Benjamini and Yekutieli, 2001; Curran-Everett, 2000; Thissen et al., 2002), with the likelihood of acceptance of at least one falsely significant result set at $p < 0.05$.

Using data from the surveys, we analyzed how students' feelings towards changing their study methods related to their likelihood of changing the percentage of study time spent on a given study method. To address this, we compared student-reported reluctance to changing study methods at the start of the course ("I am reluctant to change my study habits, because they have worked very well for me so far.") to the actual change between the start of the course and the end of the course. Only students who were confirmed to have responded at both timepoints and fully completed both relevant sections of the survey were included in this analysis. After applying this filter, there were fewer than three responses for the categories ‘Strongly Agree’ and ‘Strongly Disagree,’ and these responses were combined with ‘Agree’ and ‘Disagree,’ respectively. Using these combined groups of ‘Agree,’ ‘Neither Agree nor Disagree,’ and ‘Disagree,’ we ran an ANOVA on the amount of change in study methods for each category; this was done for each cohort for drawing, rereading, active methods, and passive methods.

2.3.2 Research question 2: the relationship between usage of sketching by students and performance

From the interviews, we assessed how students applied their study methods when answering questions on exams, as well as whether answers that were studied and answered using drawing differed in quality (based on exam point scores) from those studied and answered using other methods. In order to assess which method a student used to study or solve problems, we coded the interviews based on a series of questions (modified from Hauge, 2019; see Appendix A2 “Interview Coding Questions” for a copy of the coding rubric), including whether they drew on those exam questions or included drawing as part of their study time related to the topics of the discussed exam questions. We used t-tests to compare question scores for students who used drawing on exam questions to question scores of students who did not use drawing on exam questions. We also tested when studying using drawing was related to scores on exam questions on that topic. We used t-tests to compare question scores of students who drew while studying for the exam questions reviewed or did not draw to study for the exam questions reviewed. Due to the timing of interviews, only the cohort with the drawing intervention was included in this analysis.

We also tested whether proportion of study time spent drawing by the end of the course was related to the final course score using a regression analysis in R. This comparison only included the 41 students who both responded to the post-survey and participated in the interviews of students with the drawing intervention. We would like to repeat this analysis with the cohort without the drawing intervention, but do not currently have data on final course scores for these students.

2.3.3 Research question 3: student motivations towards drawing as a learning tool

To address questions related to motivation, the Likert responses were converted to a numerical form (i.e. “Strongly Disagree” = 1, etc.) where higher values indicated higher cost, self-efficacy, interest, and utility. One question related to the cost construct (“I am willing to spend my time to practice sketching to learn.”) was worded in the reverse format to a typical scale, so a higher value would instead indicate a lower cost to the respondent. To account for this, we reversed the scale of this question (multiply by negative one and add the total number of Likert-scale response options, 7). Once all questions were scaled in the same direction, the responses for each student were averaged into composite scores for each construct.

These composite scores were then compared between cohorts using t-tests to determine differences in expectancy-value motivation for students with and without instruction in drawing as part of their Introductory Biology course. Due to the multiple statistical comparisons conducted on these data, we used a false discovery rate control to adjust significance (Benjamini and Hochberg, 1995; Benjamini and Yekutieli, 2001; Curran-Everett, 2000; Thissen et al., 2002), with the likelihood of acceptance of at least one falsely significant result set at $p < 0.05$.

To address whether motivation predicts student reported use of drawing as a learning tool, we performed a multiple regression in R with the four motivation constructs as predictor variables and percent study time spent drawing as the outcome variable (see Appendix A3 "Code for Analyses completed in R" for R code used in analyses).

. Prior to conducting this regression, we conducted multiple imputations using the MICE package in R (using the default setting of five imputations) to account for missing responses in the motivation items and the associated use of drawing as a study tool. Using multiple imputations created five complete datasets, each of which was analyzed individually as a multiple regression using MICE. These five versions of the results were then pooled into a single result. This same process was completed to calculate the correlations between the motivation constructs and to calculate the R-squared value.

3. Results

3.1 Research Question 1: Patterns of growth and decline in the use of sketching and other active study methods

3.1.1 Within Cohort with Drawing Intervention

In order to answer our questions relating to how students study methods changed over time, we looked at participants self-reported study time. Students who failed to fully complete this section of the survey were removed from this portion of the analysis. First, we looked at how students in the cohort with the drawing intervention changed their methods between the start of the course, the end of the course, and one semester after the course. Specifically, we focused on drawing, rereading, combined active methods, and combined passive methods. All methods are reported as the average of student-reported percent of overall study time using a given method with 95% confidence intervals (CI).

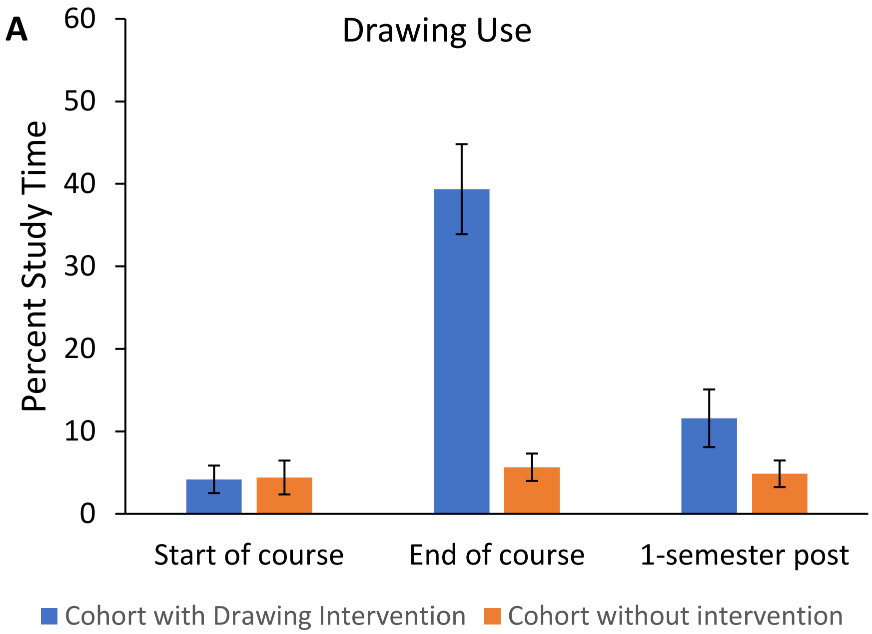
The students with the drawing intervention increased their use of drawing nearly ten-fold from the start of the course (4.18 percent ± 1.67; n=158) to the end (39.36 percent ± 5.45; n=86; $p < 0.001$). Between the end of the course and one semester later, students decreased their use of drawing to study (11.58 percent ± 3.50; n=55; $p < 0.001$) but did not return to baseline ($p < 0.001$; Figure 3A). For rereading, they began the course at 41.15 percent ± 4.10 (n=158), nearly halving time spent rereading by the end of the course (25.83 percent ± 4.94; n=86; $p = 0.001$). One semester later these students barely increased their use of rereading (28.93 percent ± 5.81; n=55; $p = 0.428$), but rereading was still significantly less than the baseline ($p = 0.003$; Figure 3B)

Active study methods including drawing made up 35.80 percent ± 3.65 of study time at the start of the course (n=158), increasing to 59.11 percent ± 5.60 (n=86; $p < 0.001$). Between the end of the course and one semester later (four months), students decreased their use of active study methods (43.58 percent ± 6.31; n=55; $p < 0.001$) but did not return to baseline ($p = 0.019$; Figure 4A). Passive study methods including rereading made up 55.46 percent ± 4.02 of study time at the start of the course (n=158), decreasing to 33.75 percent ± 5.41 (n=86; $p < 0.001$). Between the end of the course and one semester later, students increased their use of passive study methods (44.69 percent ± 5.87; n=55; $p = 0.008$) but did not return to baseline ($p = 0.015$; Figure 4B).

3.1.2 Within Cohort without Drawing Intervention

Next, we repeated this comparison for the cohort without the intervention. This group began the course with a low use of drawing (4.40 percent ± 2.05; n=77) and maintained a similar use of drawing at the end of the course (5.65 percent ± 1.66; n=66; $p = 0.287$; Figure 3A). A semester later these students reported a similar use of drawing (n= 82; 4.85 percent ± 1.62) as at the start ($p = 0.639$) and the end of the course ($p = 0.501$; Figure 3A). For rereading, they began the course at 43.31 percent ± 5.69 (n= 77), remaining roughly the same at the end of the course (46.08 percent ± 5.89; n=66; $p = 0.299$). One semester later these students maintained the same level of rereading (n= 82; 44.49 percent ± 5.94) as at the start of the course ($p = 0.187$) and end of the course ($p = 0.959$; Figure 3B).

Active study methods including drawing made up 35.51 percent ± 5.21 at the start of the course (n= 77), staying roughly the same at 30.55 percent ± 5.12 (n=66; $p = 0.331$). Between the end of the course and one semester later, students maintained their use of active study methods (n= 82; 36.02 percent ± 5.15; $p = 0.079$), at roughly the same level as baseline ($p = 0.737$; Figure 4A). Passive study methods including rereading made up 54.44 percent ± 5.81 of study time at the start of the course (n= 77), increasing to 62.05 percent ± 5.74 (n=66; $p = 0.025$). Between the end of the course and one semester later, students maintained their use of passive study methods (n= 82; 56.54 percent ± 5.73; $p = 0.090$), which was also similar to the use of passive methods at the start of the course ($p = 0.489$; Figure 4B).



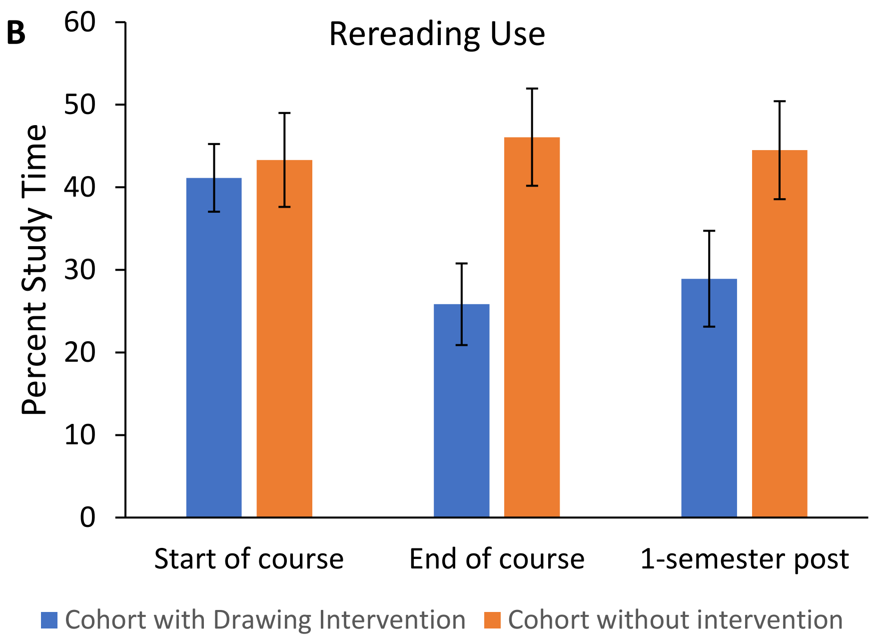


Figure 3 - **Introductory biology course incorporating drawing as a learning tool leads to increase in use of drawing and decrease in rereading.** Average student-reported percentage of study time (± 95% CI) spent drawing or redrawing (A) and rereading notes, text, or course materials (B) at the start of the course, the end of the course, and one semester after a one-semester Introductory Biology course are plotted for the cohort with an incorporated drawing intervention (blue; $n_{start} = 158$, $n_{end} = 86$, $n_{1\text{-semester}} = 55$) and the cohort without a

drawing intervention (orange; $n_{start} = 77$, $n_{end} = 66$, $n_{1\text{-}semester} = 82$). T-tests comparing the cohorts at each timepoint showed that the cohorts began the same for drawing, but differed at later timepoints for both drawing ($p_{start} = 0.886$, $p_{end} < 0.001$, $p_{1\text{-}semester} < 0.001$) and rereading ($p_{start} = 0.618$, $p_{end} < 0.001$, $p_{1\text{-}semester} < 0.001$). Additional t-tests showed that the cohort without the drawing intervention did not change between the start and end of the course ($p_{drawing} = 0.287$, $p_{rereading} = 0.299$), the start and one semester post ($p_{drawing} = 0.639$, $p_{rereading} = 0.187$), or the end and one-semester post ($p_{drawing} = 0.501$, $p_{rereading} = 0.959$). The cohort with the drawing intervention changed significantly between the start and end of the course ($p_{drawing} < 0.001$, $p_{rereading} < 0.001$), the start and one semester post ($p_{drawing} < 0.001$, $p_{rereading} = 0.003$), but only significantly for drawing from the end and one-semester post ($p_{drawing} < 0.001$, $p_{rereading} = 0.428$). All p-values maintained significance at $p < 0.05$ following assessment using the false discovery rate control.

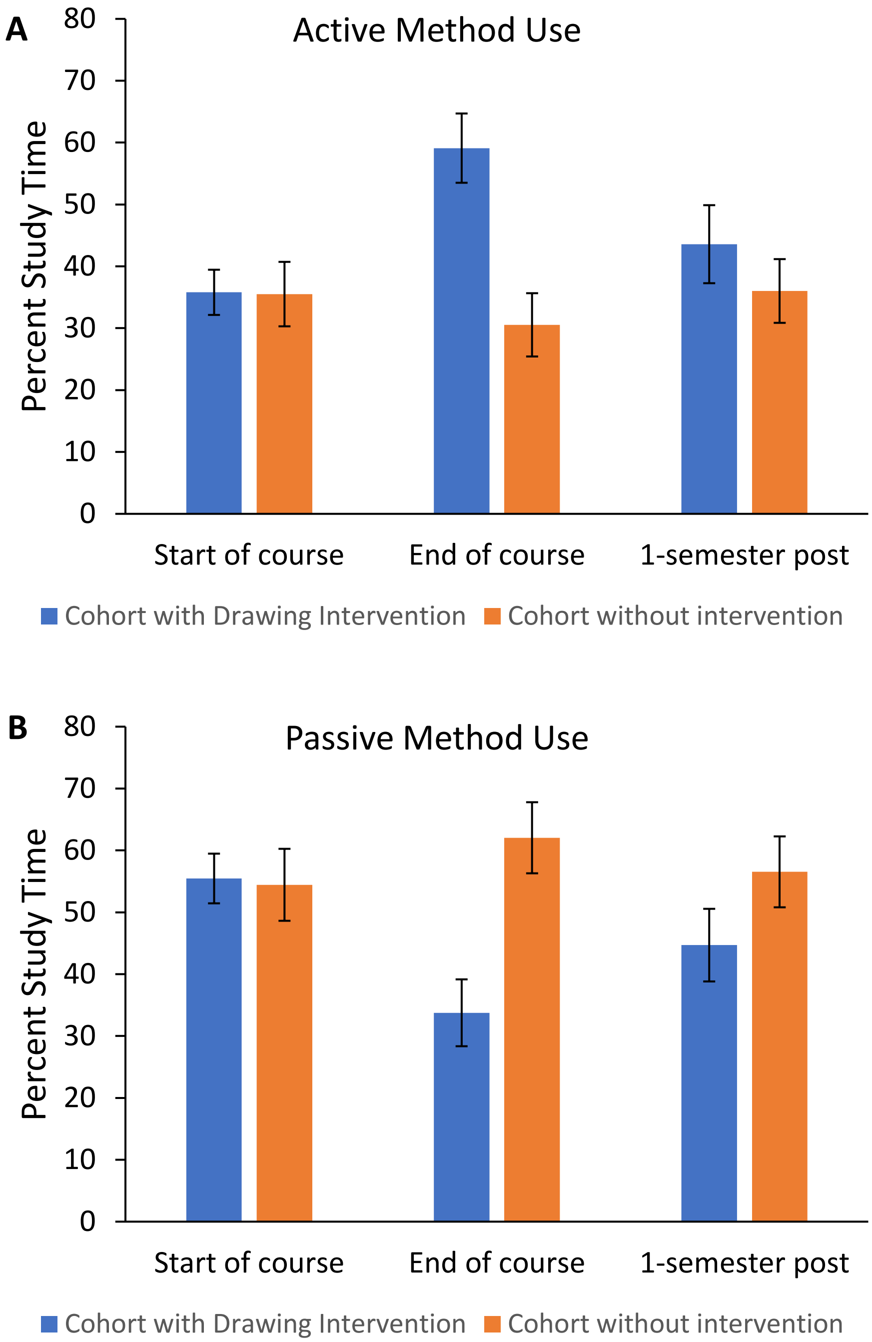


Figure 4 - **Introductory biology course incorporating drawing as a learning tool leads to increase in use of active study methods and decrease in passive study methods.** Average student-reported percentage of study time (± 95% CI) spent using active methods (A) and rereading notes, text, or course materials (B) at the start of the course, the end of the course, and one semester after a one-semester Introductory Biology course are plotted for the cohort with an incorporated drawing intervention (blue bars; n_{start} = 158, n_{end} = 86, $n_{1\text{-semester}}$ =

55) and the cohort without a drawing intervention (orange bars; $n_{start} = 77$, $n_{end} = 66$, $n_{1\text{-semester}} = 82$). Active methods include flashcards, drawing, redrawing, practice tests, self-testing, writing your own questions, chunking, retrieval practice, mind maps or concept maps, folded lists, and minute sketches. Passive methods include rereading course materials or personal notes, rewriting, and highlighting. T-tests comparing the cohorts at each timepoint showed that the cohorts began the same for, differed end of the course for both active methods, methods, but only differed one semester post-course for passive methods ($p_{start} = 0.716$, $p_{end} < 0.001$, $p_{1\text{-semester}} = 0.005$), not active methods ($p_{start} = 0.874$, $p_{end} < 0.001$, $p_{1\text{-semester}} = 0.071$). Additional t-tests showed that the cohort without the drawing intervention did not change between the start and end of the course for active methods ($p_{active} = 0.331$) but did for passive methods ($p_{passive} = 0.025$). This cohort did not change between the start and one semester post ($p_{drawing} = 0.737$, $p_{rereading} = 0.489$), or the end and one-semester post ($p_{active} = 0.079$, $p_{passive} = 0.090$). The cohort with the drawing intervention changed significantly between the start and end of the course ($p_{active} < 0.001$, $p_{passive} < 0.001$), the start and one semester post ($p_{active} = 0.019$, $p_{passive} = 0.015$), and the end and one-semester post ($p_{active} < 0.001$, $p_{passive} = 0.008$). All p-values maintained significance at $p < 0.05$ following assessment using the false discovery rate control.

3.1.3 Between Cohorts at each timepoint

Finally, we compared the usage of study methods between the two cohorts at each timepoint. Both cohorts began their semester-long Introductory biology courses with approximately the same level of drawing use (4.18 percent ± 1.67 intervention vs 4.40 percent ± 2.05 non-intervention; $p = 0.886$). However, after completing their courses, the cohort with the drawing intervention used drawing while studying nearly seven times as much as the cohort without the drawing intervention (39.36 percent ± 5.45 intervention vs 5.65 percent ± 1.66 non-intervention; $p < 0.001$). One semester after their courses, the cohort with the drawing intervention continued to use drawing more than the cohort without the drawing intervention (11.58 percent ± 3.50 intervention vs 4.85 percent ± 1.62 non-intervention; $p < 0.001$; Figure 3A). For rereading, both cohorts began their semester-long courses with approximately the same level of rereading use (41.15 percent ± 4.10 intervention vs 43.31 percent ± 5.69 non-intervention; $p = 0.618$). However, after completing their courses, the cohort with the drawing intervention used rereading while studying nearly half as much as the cohort without the drawing intervention (25.83 percent ± 4.94 intervention vs 46.08 percent ± 5.89 non-intervention; $p < 0.001$). One semester after the courses, the cohort with the drawing intervention continued to use rereading less than the cohort without the drawing intervention (28.93 percent ± 5.81 intervention vs 44.49 percent ± 5.94 non-intervention; $p < 0.001$; Figure 3B).

Use of active methods including drawing for both cohorts at the start of their courses was approximately the same (35.80 percent ± 3.65 intervention vs

35.51 percent ± 5.21 non-intervention; $p = 0.874$). However, after completing their courses, the cohort with the drawing intervention used active methods while studying more than the cohort without the drawing intervention (59.11 percent ± 5.60 intervention vs 30.55 percent ± 5.12 non-intervention; $p < 0.001$). One semester after the courses, there was no statistically significant difference in self-reported use of active study methods between the intervention group and the comparison group (43.58 percent ± 6.31 intervention vs 36.02 percent ± 5.15 non-intervention; $p = 0.071$; Figure 4A). For passive methods, both cohorts began their semester-long courses with approximately the same level of passive methods (55.46 percent ± 4.02 intervention vs 55.44 percent ± 5.81 non-intervention; $p = 0.716$). However, after completing their courses, the cohort with the drawing intervention used passive methods while studying nearly half as much as the cohort without the drawing intervention (33.75 percent ± 5.41 intervention vs 62.05 percent ± 5.74 non-intervention; $p < 0.001$). One semester after the courses, the cohort with the drawing intervention continued to use passive methods less than the cohort without the drawing intervention (44.69 percent ± 5.87 intervention vs 56.54 percent ± 5.73 non-intervention; $p = 0.005$; Figure 4B).

3.2 Research question 2: the relationship between usage of sketching by students and performance

3.2.1 Relationship between reluctance to change and changes in study methods

We observed how students' feelings towards changing their study methods relates to their likelihood of changing by comparing student-reported reluctance to changing study methods at the start of the course to the actual change between the start of the course and the end of the course with an ANOVA. This comparison was repeated for each cohort for drawing, rereading, active methods, and passive methods.

For the cohort with the drawing intervention there were no significant differences for drawing, rereading, or passive methods ($p = 0.59$, 0.26, and 0.20 respectively; Table 1A). However, there was a difference between the groups for active methods ($p = 0.02$). A Tukey’s HSD post-test showed these differences to be between ‘Agree’ and ‘Disagree’, as well as between ‘Agree’ and ‘Neither Agree nor Disagree’. Even with this correction for multiple comparisons at $p < 0.05$, the significance of this difference remained. We repeated this comparison for the cohort without the drawing intervention, finding no significant differences for any of the study methods, drawing, rereading, active methods, or passive methods ($p = 0.21$, 0.54, 0.32, 0.86 respectively; Table 1B).

Table 1 - **Student-reported reluctance to changing study habits rarely affects actual likelihood to change.** Students responded on a 5-point Likert scale to the question "I am reluctant to change my study habits, because they have worked very well for me so far." Due to the low number of responses of 'Strongly agree' and 'Strongly Disagree', these were grouped for analysis with 'Agree' and 'Disagree' respectively. ANOVAs were run comparing student reluctance level to changes in the use of study methods (drawing, rereading, active methods, and passive methods). Mean change from the start of the course to the end of the course are reported for each Likert response, along with the associated p-values, for each group's ANOVA for the cohort with a drawing intervention (A) and the cohort without the drawing intervention (B). Asterisks (*) indicate differences significant at $p < 0.05$ following assessment using the false discovery rate control.

Reluctant to Change Study Methods				
A - With Intervention	**Agree**	**Neutral**	**Disagree**	**p-value**
N	12	20	22	--
Mean Draw	31.75	40.02	40.01	0.59
Mean Reread	-5.62	-18.83	-21.44	0.26
Mean Active	7.71	32.88	34.48	*** 0.02**
Mean Passive	-10.87	-27.53	-28.28	0.20

B - No Intervention	**Agree**	**Neutral**	**Disagree**	**p-value**
N	7	6	3	--
Mean Draw	-1.43	6.83	5.33	0.21
Mean Reread	10.71	9.67	-5.33	0.54
Mean Active	2.43	-8.83	-12.00	0.32
Mean Passive	11.43	12.17	4.67	0.86

3.2.2 Relationship between drawing use and academic performance

We compared the student-reported percent of study time spent drawing or redrawing to the final grade of students in the class. Since the identity was known for only the subset of students interviewed, this analysis included 41 students. We performed a linear regression in R of the final grade (percent) in the course on self-reported time spent drawing or redrawing while studying at the end of the course with the drawing intervention. This regression showed a small positive relationship between drawing while studying and final course grade ($y = 0.1634x + 69.7418$; $R^2 = 0.0879$; $p = 0.034$; Figure 5: we intend to rerun this regression using resampling procedures such as bootstrapping because of the variance structure in this analysis). We also ran a t-test between those who drew on an exam question discussed in interviews and the participant's percent score on those questions. There was no significant difference between these groups ($p = 0.977$).

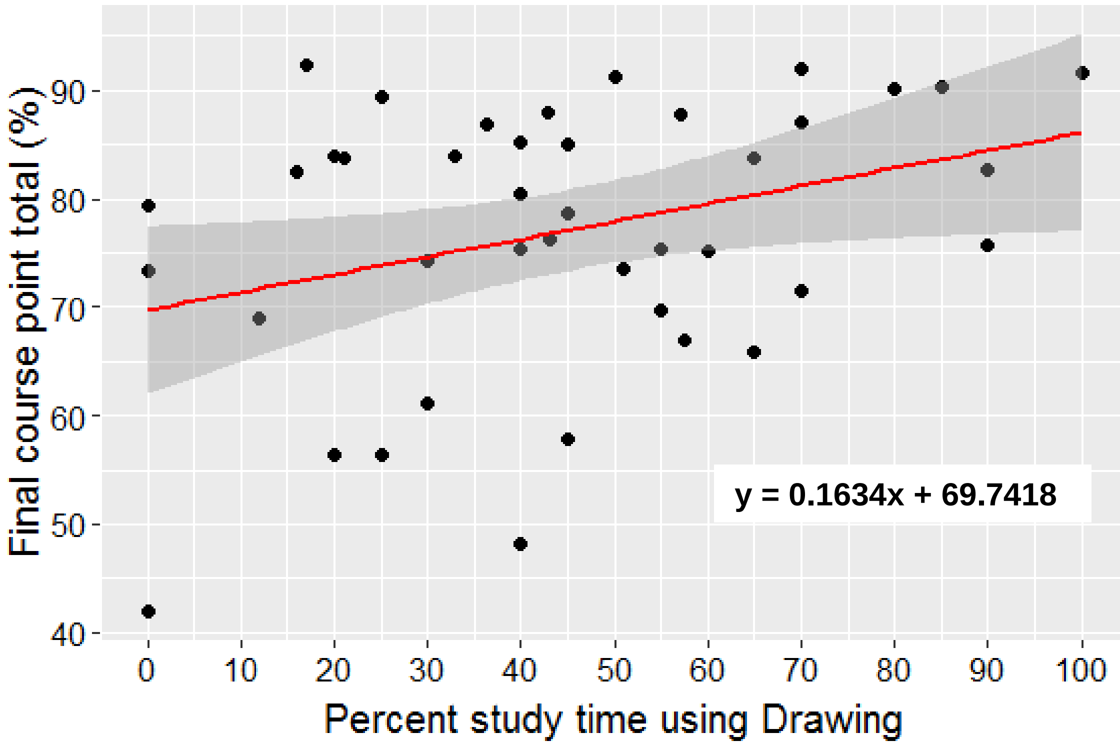


Figure 5 – **Time spent drawing while studying has a predicts final grade in course.** We performed a linear regression of the final grade (percent) in the course on self-reported time spent drawing or redrawing while studying at the end of the course with the drawing intervention (n=41; R^2 = 0.0879; p= 0.034). For every percent increase in study time using drawing, the students see a 0.1634 percent increase in final course point value.

3.3 Research question 3: student motivations towards drawing as a learning tool

3.3.1 Expectancy-value motivation towards drawing as a learning tool

Table 2 – **Pearson correlations between motivation constructs and drawing use.** “Drawing Use” is the student reported percent of study time spent drawing.

Variable	Mean	SD	1	2	3	4
1. Self-Efficacy	3.137	0.858				
2. Cost	3.737	1.210	-0.452***			
3. Interest	4.714	1.262	0.314***	-0.646***		
4. Utility	5.166	1.199	0.439***	-0.697***	0.623***	
5. Drawing Use	7.555	10.660	0.371***	-0.563***	0.451***	0.466***

*p < 0.05
**p < 0.01
***p < 0.001

One-semester after taking Introductory Biology with an integrated intervention on drawing to learn, students were asked to answer a series of questions about interest, utility, cost, and self-efficacy in regard to drawing as a learning tool. The correlations between all motivation constructs and between the constructs and the use of drawing during study time were found to be statistically significant ($p < 0.001$ for all comparisons; Table 2). Students reported on average above neutral scores of interest and utility (4.71 ± 0.37 and 5.15 ± 0.33 respectively, where neutral = 4; Figure 6A and C), as well as self-efficacy (3.39 ±

0.20, where neutral = 3; Figure 6D). They also reported a below average score for cost (3.49 ± 0.35, where neutral = 4; Figure 6B).

This analysis was repeated for the cohort without the drawing intervention. These students reported on average above neutral scores of interest and utility (4.69 ± 0.27 and 5.15 ± 0.26 respectively, where neutral = 4; Figure 6A and C), but below neutral scores of self-efficacy (2.95 ± 0.19, where neutral = 3; Figure 6D). They also reported a slightly below average score for cost (3.89 ± 0.25, where neutral = 4; Figure 6B). When comparing these two cohorts, the group with the drawing intervention had statistically significantly higher self-efficacy than the cohort without the drawing intervention ($p = 0.003$). None of the other comparisons were significantly different ($p_{interest} = 0.953$; $p_{cost} = 0.066$; $p_{utility} = 0.999$).

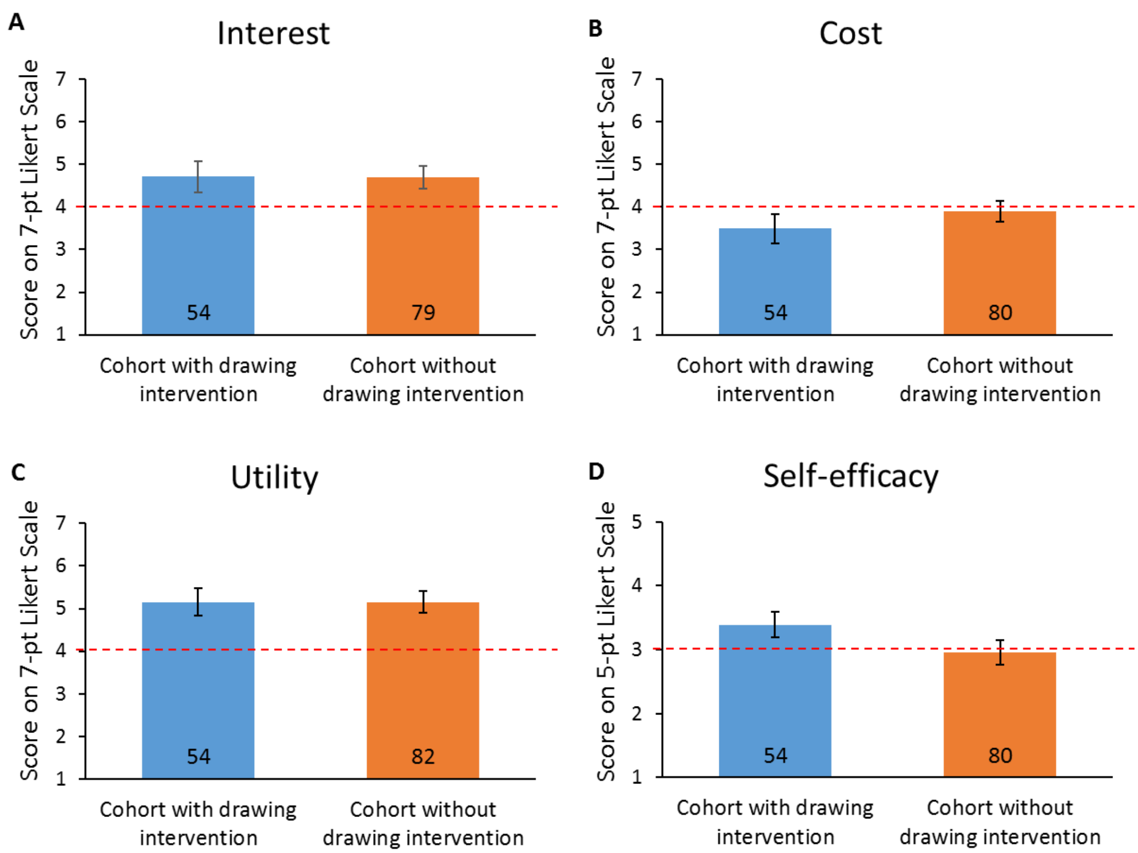


Figure 6 - **Students' expectancy-value motivation towards drawing as a learning tool.** One semester after a course in Introductory Biology students from a cohort with instruction in drawing (blue bars) and without instruction in drawing (orange bars) were asked to respond to a series of questions regarding interest, utility, cost, and self-efficacy towards drawing as a learning tool. Bars reaching above the neutral line (indicated by a red dashed line on the graphs) for interest (Panel A), utility (Panel C), and self-efficacy (Panel D) suggest positive feelings towards drawing as a learning tool, while for cost (Panel B) this suggests negative feelings. Sample size is indicated by the numbers in the base of each bar (± 95% CI). There was a significant difference between the cohorts at $p < 0.05$ following assessment using the false discovery rate control for self-efficacy ($p = 0.003$), but none of the other comparisons ($p_{interest} = 0.953$; $p_{cost} = 0.066$; $p_{utility} = 0.999$).

3.3.2 Relationship between expectancy-value motivation and drawing usage

Since the literature suggests students struggle to adopt drawing as a learning skill, we wanted to determine factors that affect their likelihood of adopting the method. We considered four types of motivation covered by the expectancy value theory: utility value, cost value, interest value, and self-efficacy. Each of these motivations were measured though student-reported feelings towards drawing as a learning tool. We then tested for whether each of these motivation types predicts student-reported use of drawing using a multiple regression in R.

We ran a multiple regression of proportion of study time spent drawing on the four motivation constructs, utility value, cost value, interest value, and self-efficacy (Table 3). When comparing individual students' motivation rankings to their use of drawing as a study method, only cost significantly predicted drawing use tool (y = 0.6293 Utility + 1.6478 Self-efficacy + 1.0465 Interest – 3.2885 Cost + 6.4916); n = 137; $R^2 = 0.3481$; $p_{cost} < 0.001$, $p_{interest} = 0.207$, $p_{utility} = 0.504$, $p_{self\text{-}efficacy} = 0.102$). This relationship indicated that students who ranked the cost of drawing as a learning tool lower were more likely to use higher proportions of their study time drawing (shown as a single regression in Figure 7: we intend to rerun this regression using resampling procedures such as bootstrapping and as a Poisson regression because of the variance structure and count-style, zero-inflated data involved in this analysis).

Table 3 – **Multiple Regression of Percent Time Drawing to Study on Motivation Constructs (n = 137)**

Parameter	Estimate	SE	t-value	p-value
Intercept	6.492	8.660	0.749	0.455
Self-Efficacy	1.648	0.999	1.649	0.102
Cost	-3.288	0.958	-3.431	0.0008
Interest	1.046	0.826	1.267	0.207
Utility	0.629	0.939	0.670	0.504

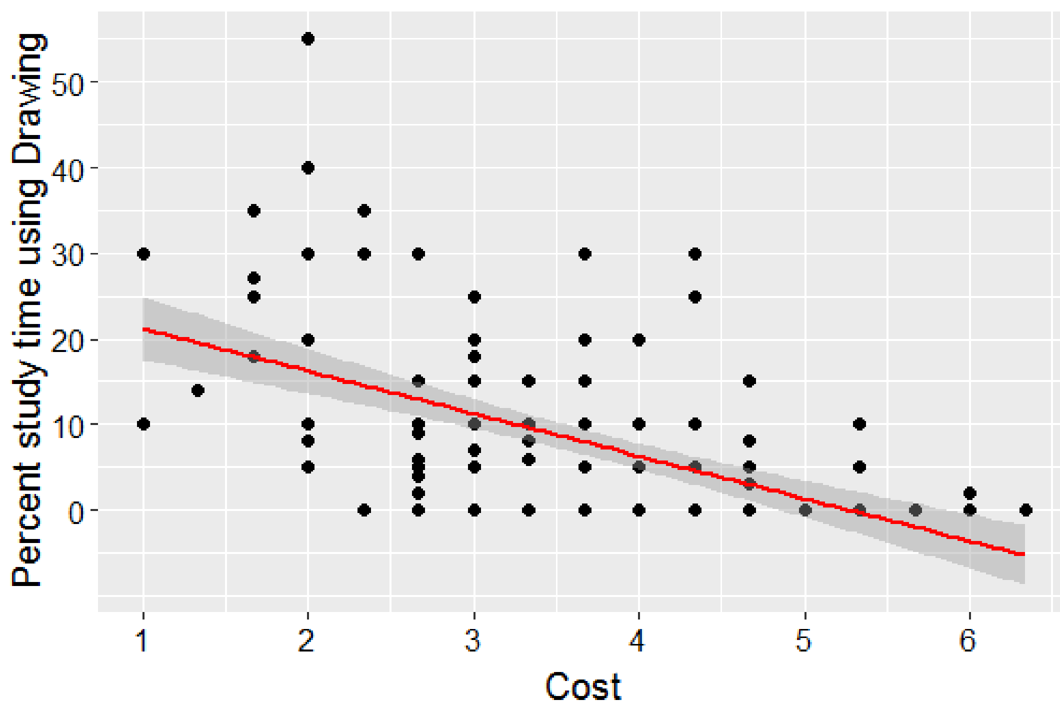


Figure 7 - **Students' self-reported cost value motivation towards drawing as a learning tool predicts use of drawing as a study tool.** One semester after a course in Introductory Biology students from both the cohort with instruction in drawing and students without instruction in drawing were asked to respond to a series of questions regarding interest, utility, cost, and self-efficacy towards drawing as a learning tool. Using a multiple regression in R, we analyzed whether each motivation construct predicts student use of drawing as a study tool, finding only cost predicted use of drawing as a study tool (y = 0.6293Utility + 1.6478Self-efficacy + 1.0465Interest – 3.2885Cost + 6.4916; n = 137; R^2 = 0.3481; $p_{cost} < 0.001$, $p_{interest} = 0.207$, $p_{utility} = 0.504$, $p_{self\text{-}efficacy} = 0.102$). Shown is a single regression of the only significant effect: cost (n = 137; R^2 = 0.3164; $p < 0.001$).

4. Discussion

4.1 Research Question 1: Patterns of growth and decline in the use of sketching and other active study methods

Initial results suggest that this guided practice method of integrating instruction on drawing as a study tool increased student usage of this tool both during the course and one semester later. The large size of the increase during the course was expected since students were required to use drawing as an integrated portion of their homework before each class session. This is consistent with the results found with a shorter intervention of this type (Heideman et al., 2017), suggesting that drawing taught as a short intervention or as a semester-long integrated intervention may be effective for short-term increases in drawing use.

Although the intervention was effective in the short-term, it is more important to ask about effectiveness for the long-term. Professors have limited time for instruction, and it would be inefficient for every instructor to teach drawing skills, even if all instructors were motivated to do so. For our approach to teach drawing in biology to be worthwhile for instructors and students, we expect that long-term retention of motivation to draw may be required. Our results show that a semester after the course with integrated drawing, students continued to use drawing more than students from the non-drawing course. However, the amount of drawing to study dropped significantly from the use of drawing at the end the course. This decrease in drawing use one semester after the course was understandable, as students were no longer required to draw or sketch as homework, which by default would decrease their use of this method. This drop is also consistent with an unpublished follow-up on the Heideman et al (2017) study. In that study the amount of drawing used by students had returned to the baseline by 2.5 years after the intervention. In our study, we do not yet have data on longer term outcomes.

These same patterns were seen in the cohort with the drawing intervention in relation to combined active methods. Since active methods are typically more effective (Dunlosky et al., 2013; Fu and Gray, 2004), this pattern of change suggests that the form of intervention focused on improved studying, as was taught to the cohort with instruction on drawing, may help students to shift their study habits to active methods as the students were trying to be more effective learners. Importantly, these students retained these more effective active methods a semester after they were asked to focus on their study habits. Integrating instruction on improving study habits, including drawing, may promote increased use of more effective study skills and a better understanding of course content.

The students with instruction on drawing for biology also decreased their use of rereading and other passive study methods. This means that rather than simply changing their overall study time (i.e. by adding both more active and more passive methods), students were changing the proportion of study time spent on passive methods, devoting a greater proportion of study to be time spent on the active methods. This suggests that these students, whether intentionally or not, were able to shift their habits to methods typically found more effective.

The cohort without the drawing intervention showed almost no changes in study habits across the year of data collection. Although we expected there to be little change in use of drawing, we were surprised by the general lack of change over the first semester. We had expected a shift towards more active methods because many of the students in this research were early in their academic careers, and thus they might alter their habits slightly to adjust to the higher demands of university compared to high school. However, we saw the reverse. In the cohort without instruction on more active study methods, students slightly increased their use of passive methods. This suggests that self-efficacy may be important because students likely already feel confident in their ability to apply these passive methods to coursework. Passive methods are generally less effective (Dunlosky et al., 2013; Fu and Gray, 2004) and therefore may lead to smaller learning gains. This is a vital point, consistent with the literature (Pressley et al., 1989; Donovan and Bransford, 2005; Fu and Gray, 2004), as it suggests that students fail to adopt more effective study methods even when the demands on their time and memory presumably increase during college.

This lack of increase in active methods may suggest that students have lower self-efficacy towards active methods, including drawing. Students may not be explicitly taught active study methods and habits, which typically lowers self-efficacy (Schunk, 1989). This was supported by our results showing that students in the cohort without the instruction on drawing also reported statistically significantly lower self-efficacy for drawing as learning tool.

One semester after the Introductory Biology course, there was no statistically significant difference in drawing or active methods for students without instruction on drawing as a learning tool, but these students used statistically equivalent amounts of active methods when compared to the drawing cohort. This appears to be due to a slight increase in active methods by the non-drawing cohort and a statistically significant decrease in active methods by the cohort with the drawing intervention. This could be due to chance, but it may also suggest that the cohort without the drawing intervention started to see the small increase we had expected in active methods after a year in university-level courses.

4.2 Research question 2: the relationship between usage of sketching by students and performance

Our results showed mixed results on the difference in academic performance between students who used drawing and those who did not use drawing. First, we considered whether percent of study time spent drawing might predict a student's final grade in the course. We found a small positive relationship between drawing use during studying and academic performance (Figure 5). This suggests that students who spent a higher proportion of time drawing to study might perform better in BIOL 203, with a between small and moderate effect size of 0.09 (reported as R^2: Cohen, 1992; Cohen, 1988).

The interviewed students included in this comparison had a range of grades, from failing to high A's, but were significantly different from those not interviewed. Students interviewed had on average a higher grade than the students who were not interviewed (76.68 percent and 72.40 percent respectively; $p = 0.047$). One possibility is that higher achieving students may feel less pressured by their coursework and therefore more likely to accept offers to participate in interviews. This seems unlikely considering the broad spread of interviewee grades. A second possibility is that interviewed students were influenced by the research, leading them to think more metacognitively about their study habits and problem-solving approaches, which in turn led to higher grades. This is important to consider because it may indicate that interviewee results are not fully representative of the population involved in this study.

Higher proportions of drawing while studying predicted higher course grades. But when we compared scores on questions we reviewed, students who used or did not use drawing did not perform differently on those questions. This is inconsistent with past studies done by our group (Hauge, 2018), and other studies on drawing and memory (i.e. Wammes et al., 2016), which have suggested that using drawing increases recall and academic performance. A potential explanation is that drawing as a learning tool is challenging to learn and takes time to learn (Quillin and Thomas, 2015; Van Meter and Garner, 2005). Given the short time period these students had to practice this skill before they solved the exam problems in this course, we would not necessarily expect drawing to predict performance during the first semester students were using this skill. Additionally, lower cost with higher utility and higher interest, as we observed a semester after the course, can lead to higher performance (Rosenzweig et al. 2020; Hulleman and Harackiewicz, 2009), but the short period students spent practicing this skill before the exams we reviewed may not have been enough time for students to feel comfortable with drawing as a learning and problem-solving tool. Alternatively, while drawing is a valuable skill for students to learn (Quillin and Thomas, 2015; Van Meter and Garner, 2005), many exam questions may not require processing for which drawing is needed in, for example, model-based reasoning. Drawing on an exam may not be needed if a question asks students to describe memorized text or figures, or explain a well-recalled system.

Finally, this difference in results may have occurred because all students in the course with the drawing intervention were required to draw as part of their coursework. Many of the exam topics had been drawn in homework assignments, even though students might not have drawn additional times while studying. Thus, even students who claimed not to draw while studying had regularly spent time drawing for homework. This means that students who drew by choice because they felt it was a useful study or problem-solving method were difficult to distinguish from students who drew because they felt it was a requirement of the course. Especially early in the course, this may have hindered students who were not yet comfortable with drawing as a learning tool. Once we complete and analyze interviews with students from the cohort without the drawing intervention, we will be able to compare the cohorts. There may be differences relatable to practice sketching, but there were many other differences between the courses.

Additionally, exam questions were grouped for comparison, rather than comparing scores on individual questions. In other words, rather than comparing drawing use to the score of a particular question on a particular exam, drawing use was compared to question scores regardless of the question type or topic. This was done with the expectation that any effect of drawing on exam scores would not be dependent on question type or topic, but rather would be universal across a students' exams. However, in interviews and surveys some students mentioned that they found drawing more useful for certain topics. For example, Participant 1 stated they "*...found that when remembering structures this method is extremely helpful, which is why it is amazing for biology. I do not use this method for chemistry because a lot of that is math, which I find easier to learn by doing problems.*" Participant 35 found that "*sketching is best used for understanding structures or processes*", a point brought up by several other students. This suggests that a relationship with exam question scores may only exist when considering topics with specifics characteristics, such as when a question considers multi-step processes or detailed structures.

4.3 Research question 3: student motivations towards drawing as a learning tool

Before educators can address decreasing student resistance to using new skills, we must first know why there is resistance. The motivation aspect of this research aims to determine the reasons students use or avoid the use of drawing as a learning and problem-solving tool in biology. After the introductory course with instruction on drawing, student expectancy-value scores show that on average students from the course with the drawing intervention found drawing as a tool to be useful and interesting, and had high self-efficacy towards their ability to use this tool (Figure 5A,C, and D). Surprisingly, they also found drawing as a learning tool to not be highly costly in regard to time and effort (Figure 5B). When students report scores of self-efficacy, utility, and interest that are above average or neutral, this suggests that they feel positively towards drawing as a learning tool. If students reported scores of cost above average, this would suggest that they find this tool high in cost, meaning drawing to learn requires high time commitment and effort. The literature suggests that drawing as a learning tool is difficult to learn and apply, which would predict that the cost value towards this method should be high. The fact that our students on average ranked cost as neutral suggests that the instruction in drawing may have lowered the cost for the students. This instruction included guidance on how to keep drawings simple and quick to reproduce, examples, demonstrations, and feedback that may have helped students feel that less effort and time was needed to use drawing as a learning tool.

As in the cohort with the drawing intervention, the students without the instruction on drawing had above average interest and utility, and below average cost. However, they also had statistically lower self-efficacy towards drawing as a learning tool a semester after their Introductory Biology course. This suggests that the guided practice method of teaching drawing used in the course with the intervention may have increased student self-efficacy. This was unsurprising because drawing was built into the course and exams were structured to allow, and more importantly not disadvantage, the use of drawing by students. Based on the literature, this suggests students would also be more likely to continue to use drawing as a learning tool because “the higher the sense of efficacy, the greater the effort, persistence, and resilience” (Pajares, 1996).

Students in both cohorts ranked cost below neutral, and interest and utility above neutral, but self-efficacy statistically significantly differed between the cohorts. The cohort with instruction in drawing reported higher self-efficacy than the students without instruction on drawing. This suggests that self-efficacy may be an important construct to focus on, since the instruction may have increased students’ confidence in their ability to draw as a learning tool. However, the results of our multiple regression add another level of complexity to the story. When comparing individual students’ motivation rankings to their use of drawing as a study method, only cost predicted drawing use (Table 3 and Figure 7), with a large effect size of 0.35 (reported as R^2: Cohen, 1992; Cohen, 1988). Students who reported low cost used drawing for a higher proportion of study time, while those who reported high cost were less likely to use drawing to study. This means that despite interest and utility being high, and self-efficacy differing between the cohorts, none of these three predict students use of drawing as a learning tool. These results suggest that educators teaching drawing as a learning or problem-solving tool should focus on decreasing the perceived cost of using this tool. By lowering the perceived cost, we lower the resistance towards adopting this tool, encouraging biology students to learn useful skills earlier in their careers.

This feeling of high cost was reflected by some of the comments made by participants. For example, one student (Participant N60) said drawing was "*time consuming"*, adding that "*If I draw it wrongly, I remember it wrongly*." Although studying incorrect material leading to faulty understanding is arguably a problem with any study method, students often noticed the cause of their errors while using drawing. This was likely because students were able to compare their own drawings to corrected versions and therefore see exactly where their mistake fell. Participant 21 mentioned that with sketches "*I get to talk through the process, through retrieval. And it makes it easier to spot where a mistake is made.*" Although noticing that errors in studying can lead to later errors in understanding may feel costly to the students, it also suggests that drawing helps them self-correct.

Many students complained during interviews and in surveys that drawing has a high cost, taking a lot of time and effort before feeling comfortable with drawing. Since cost predicted use of drawing, it may be important for instructors to focus on decreasing the perceived cost of drawing to learn for students. Students were not being lazy, they were simply reacting to the effort it takes to learn a new, challenging skill that typically has not been integrated into their lives previously. Based on conversations with students, educators, and other research lab members, we have outlined a few ways to address this issue. First, increasing the amount of feedback to students could provide them with more guidance, decreasing the time and effort (cost) students have to commit to self-correcting and adjusting their drawings. Secondly, scaffolding the difficulty of sketches may help decrease perceived cost to the students. Providing students with simpler drawing assignments (for example including a checklist of items to include or pieces of a sketch that they have to piece together like a puzzle) until they better understand the process of creating sketches as learning tools may make the task feel more manageable. Lastly, it may help to allow students to create sketches during short stretches of class time, then compare these to their fellow classmates before editing sketches to make them more accurate or understandable. This would give real-time feedback and chances to create and critique sketches separately from the instructor. This student interaction may decrease cost as students would be forced to draw quickly (rather than focusing on sketches being "pretty") and might come to understand that drawings can be altered and improved easily.

In addition to addressing cost to students, it may be important to address cost to the instructors. Teaching a new skill in addition to the required course content takes time, planning, and effort from the instructor. This can be decreased by making the drawing assignments low stakes (upon effort, rather than correctness) and giving only one piece of feedback per student per assignment. Additionally, with our approach most of this skill can be taught outside of the classroom as part of homework assignments. Instructors can have students watch pre-made videos on how to draw, as well as examples of topics relevant to class (i.e. Paul Heideman, 2018; with many other drawing videos included on the same channel). While individual videos may require significant effort, once suitable videos are made they can be reused indefinitely by any instructor. Many of the videos assigned to students in this study are freely available on YouTube, and others were created by our research team for other instructors. Some are being used by instructors on our campus or elsewhere by instructors wishing to apply drawing as a learning tool. As the resource library of video sketches increases, the cost to the instructor continues to decrease.

4.4 Future work

This thesis only addresses the first two years of this project. Future work on this 5-year project will allow us to assess the longitudinal changes in the students of both cohorts throughout their undergraduate careers. Additionally, we will extend our analyses to more timepoints and additional comparisons between the cohorts, including effect sizes for key comparisons. For example, analysis of interviews will allow us to look at retention of biological processes over time in both cohorts of students. This will be done by scoring the concept or process questions for correctness and complexity, then comparing this result to how that student studied and whether or not they received the drawing course. We also plan to compare students within the course with the drawing intervention to students without the drawing intervention in terms of whether answers that were studied and answered using drawing differed in quality from those not using drawing. These students will be reassessed yearly, allowing for long-term understanding of how biology students change throughout their undergraduate education. With this understanding, we aim to better provide educators and students with tools to become successful in their careers.

5. Literature Cited

Baddeley, A. (1992). Working Memory. *Science, 255*(5044), 556-559.

Bandura, A. (1997). *Self-efficacy: The Exercise of Control*. New York: W. H. Freeman, 1997.

Benjamini, Y., & Hochsberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological), 57*(1), 289-300.

Berry, J. M. (1987). A self-efficacy model of memory performance. Paper presented at the meeting of the American Psychological Association, New York.

Bielaczyc, K., Pirolli, P., & Brown, A. (1995). Training in Self-Explanation and Self-Regulation Strategies: Investigating the Effects of Knowledge Acquisition Activities on Problem Solving. *Cognition and Instruction, 13*(2), 221-252.

Bouffard-Bouchard, T. (1990). Influence of Self-Efficacy on Performance in a Cognitive Task. *Journal of Social Psychology, 130*(3), 353-363.

Chandler, P., & Sweller, J. (1991). Cognitive Load Theory and the Format of Instruction. *Cognition and Instruction, 8*(4), 293-332.

Clark, J., & Paivio, A. (1991). Dual Coding Theory and Education. *Educational Psychology Review, 3*(3), 149-210.

Cohen J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2ed., pp 413-314). Hillsdale, NJ: Laurence Erlbaum Associates.

Cohen, J (1992). "A power primer". *Psychological Bulletin*. 112 (1): 155–159

Curran-Everett, D. (2000). Multiple comparisons: philosophies and illustrations. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 279*(1), 1-8.

Dauer, J., & Long, T. (2015). Long-term conceptual retrieval by college biology majors following model-based instruction. *Journal of Research in Science Teaching, 52*(8), 1188-1206.

Donovan, M., & Bransford, J. (Eds.). (2005). *How Students Learn: Science in the Classroom.* Washington, D.C.: National Academies Press.

Dunlosky, J., Rawson, K., Marsh, E., Nathan, M., & Willingham, D. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, Supplement, 14(1)*, 4-58.

Dye, K., & Stanton, J. (2017). Metacognition in upper-division biology students: Awareness does not always lead to control. *CBE Life Sciences Education, 16*(2), 1-14.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behaviour: An introduction to theory and research.* Reading, MA: Addison-Wesley.

Fu, W., & Gray, W. (2004). Resolving the paradox of the active user: Stable subuptional performance in interactive tasks. *Cognitive Science, 28*(6), 901-935.

Gilbert, J. (2005). Visualization: A Metacognitive Skill in Science and Science Education. In J. Gilbert, & J. Gilbert (Ed.), *Visualization in Science Education* (pp. 9-27). Dordrecht: Springer.

Hauge, E. (2018). Sketching Science: How Students Use a Drawing-to-learn Based Method to Learn and Solve Problems in Undergraduate Biology. *Honors Thesis from William & Mary.*

Heideman, P. D., Flores K., Sevier L. & Trouton K. (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*(28), 1-13.

Heideman, P. D., Wilkin, L. & Williams, P. (2019). Manuscript in preparation.

Hidi, S. & Renninger K. A. (2006) The Four-Phase Model of Interest Development. *Educational Psychologist, 41(2)*, 111-127.

Hmelo-Silver, C., & Pfeffer, M. (2004). Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. *Cognitive Science, 28*(1), 127-138.

Hulleman, C. S., Godes, O., Hendricks, B. L., & Harackiewicz, J. M. (2010). Enhancing interest and performance with a utility value intervention. *Journal of Educational Psychology, 102*(4), 880-895.

Hulleman, C., & Harackiewicz, J. (2009). Promoting Interest and Performance in High School Science Classes. *Science, 326*(5958), 1410-1412.

Lent, R., Brown, S., & Larkin, K. (1984). Relation of Self-Efficacy Expectations to Academic Achievement and Persistence. *Journal of Counseling Psychology, 31*(3), 366-362.

Lent, R., Brown, S., & Larkin, K. (1986). Self-Efficacy in the Prediction of Academic Performance and Perceived Career Options. *Journal of Counseling Psychology, 33*(3), 265-269.

Mathewson, J. (1999). Visual-Spatial Thinking: An Aspect of Science Overlooked by Educators. *Science Education, 83*(1), 33-54.

Mayer, R. E. (2003). The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction, 13*(2), 125-139.

Mayer, R. E. Cognitive Theory of Multimedia Learning. (2005). In Richard E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (1 ed., pp. 31-48). Santa Barbara: Cambridge University Press.

Mayer, R. (2009). *Multimedia Learning* (2 ed.). New York: Cambridge University Press.

Mayer, R. E., Griffith, E., Jurkowich, I., & Rothman, D. (2008). Increased Interestingness of Extraneous Details in a Multimedia Science Presentation Leads to Decreased Learning. *Journal of Experimental Psychology: Applied, 14*(4), 329-339.

Miller, G. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Prcessing Information. *The Psychological Review, 63*(2), 81-97.

Multon, K., Brown, S., & Lent, R. (1991). Relation of Self-Efficacy Beliefs to Academic Outcomes: A Meta-Analytic Investigation. *Journal of Counseling Psychology, 38*(1), 30-38.

Paivio, A. (1990). *Mental Representations: A dual coding approach.* New York: Oxford University Press.

Paivio, A. (2007). *Mind and Its Evolution: A Dual Coding Theoretical Approach.* Mahwah, NJ: Erlbaum.

Pajares, F. (1996). Self-Efficacy Beliefs in Academic Settings. *Review of Educational Research, 66*(4), 543-578.

Paul Heideman. (2018, July 22). *Proteins as Motors ATP synthase NEW 7 12 2018* [Video]. YouTube. https://www.youtube.com/watch?v=lbsgzI-bF1A

Pintrich, P., & Garcia, T. (1991). Student goal orientation and self-regulation in the college classroom. In P. Pintrich, T. Garcia, & M. L. Maehr (Eds.), *Advances in Motivation and Achievement* (pp. 371-402). Greenwich, CT: JAI Press.

Pintrich, P., & Schunk, D. (2002). *Motivation in Education: Theory, Research, and Applications* (2 ed.). Upper Saddle River, NJ: Prentice Hall.

Pressley, M., Goodchild, F., Fleet, J., Zajchowski, R., & Evans, E. (1989). The challenges of classroom strategy instruction. *The Elementary School Journal, 89*(3), 301-342.

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), 1-16.

Rosenzweig, E. Q., Wigfield, A., & Hulleman, C. S. (2020). More useful or not so bad? Examining the effects of utility value and cost reduction interventions in college physics. *Journal of Educational Psychology, 112*(1), 166-182.

Schunk, D.H. (1989). Self-Efficacy and Achievement Behaviors. *Educational Psychology Review, 1*(3), 173-208.

Schunk D.H. & DiBenedetto M.K. (2016). Self-efficacy Theory in Education. In K.R. Wentzel & D.B. Miele (Eds), *Handbook of Motivation at School* (2ed., pp. 34-54). New York, NY: Routledge.

Schunk, D.H. & Usher, E.L. (2012). Social Cognitive Theory and Motivation. in R.M. Ryan (Ed.), *The Oxford Handbook of Human Motivation* (pp 13-27). New York, NY: Oxford University Press

Simon, H. (1974). How Big Is a Chunk? *Science, 183*(4124), 482-488.

Sweller, J. (1999). *Instructional design in technological areas.* Camberwell, Australia: ACER Press.

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In J. Sweller, & Richard E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19 –31). New York: Cambridge University Press.

Thissen, D., Steinberg, L., & Kuang, D. (2002). Quick and Easy Implementation of the Benjamini-Hochberg Procedure for Controlling the False Positive Rate in Multiple Comparisons. *Journal of Educational and Behavioral Statistics, 7*(1), 77–83.

Usher, E. L., & Pajares, F. (2008). Sources of Self-Efficacy in School: Critical Review of the Literature and Future Directions. *Review of Educational Research, 78*(4), 751–796.

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

Wammes, J., Meade, M., & Fernandes, M. (2016). The drawing effect: Evidence for reliable and robust memory benefits in free recall. *Quarterly Journal of Experimental Psychology, 69*(9), 1752-1776.

Wigfield, A., & Eccles, J. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68-81.

Wu, S., & Rau, M. (2018). Effectiveness and efficiency of adding drawing prompts to an interactive educational technology when learning with visual representations. *Learning and Instruction, 55*, 93-104.

Appendices

A1. Consent Form Example

Dear Students,

Thank you for your interest in our research experiment on learning. If you agree to participate, you will take a short Qualtrics survey and will be asked to sign up for an "interview" session. In each session, which will last about 25 minutes (no more than 30), the experimenter will ask you to talk through your thought process for some of the exam questions. Please note that it makes no difference whether your answers were correct or not. We encourage all students to participate. Each session will be audio-recorded for later analysis by the researchers.

Before analyzing any of the materials, your name will be converted to a number code. Your specific responses will not be shared with others in association with your name or with identifying information. Only the interviewers (J. Burns, E. Arents, A. Tan, N. Lignore, E. Watson, N. Harris, M. Zamecnik) will review the audio sessions, and identifying information will be kept confidential. Audio transcripts of the interviews will be identified only by a randomly chosen project number. Transcripts will be saved in password-protected files (file names with an experimental ID number, but no identifying information). The passwords will be known only to the interviewers. After the audio recordings of interviews have been reviewed and transcribed, they will be deleted.

We will provide a copy of the final results to individual participants who request a copy. Your participation is completely optional and you may opt out of the research at any point without any consequences. You will receive compensation of $10 for the interview session and survey you register for and attend.

If you have any concerns about the research process you may contact the Chair of the Protection of Human Subjects Committee, Dr. Jennifer A. Stevens (jastev@wm.edu, 757 221 3862).

Please contact Jessica Burns (jrburns01@email.wm.edu) if you have questions. Thank you!

Paul D. Heideman
Boles-Ash Distinguished Professor of Biology, College of William & Mary

Student: I, ________________________, agree to participate in this experiment on learning. I understand that I may opt out of the study at any time without penalty.

Signature: ___________________________________

Date: _____________

THIS PROJECT WAS APPROVED as PHSC-PHSC-2019-08-05-13782-pdheid BY THE COLLEGE OF WILLIAM AND MARY PROTECTION OF HUMAN SUBJECTS COMMITTEE ON 2020-08-25 AND EXPIRES ON 2020-08-25.

A2. Interview Coding Questions

1) Participant #
2) Exam/quiz name
3) Question #
4) Question Type (MC or FR)
5) Was drawing required by the question?
6) Student score on a problem (numerical)
7) Total Possible points on a problem (numerical)
8) Did the student draw on the exam? (0 - N, 0.5 - edited provided image, 1 - Y, blank - unclear/no response)
    a) For a simple comparison of whether a student drew, 0.5 will be considered a 1. This is tracked separately in case we want to consider this difference later.
9) Did the student write on the exam beyond just labels? (0 - N, 0.5 - Math equations & Punnett squares, 1 - Y, blank - unclear/no response)
10) Was the first information thought of visual? (0 - not visual/verbal, 1 - visual, blank - unclear/no response)
11) Was the first information thought of a sketch? (0 - N, 1 - Y, blank - unclear/no response)
12) Did the student start drawing before they began writing on the exam? (0 - N, 1 - Y, blank - unclear/no response)
13) Did the student practice sketching for this topic (for studying)? (0 - N, 0.5 - only when assigned for homework, 1 - Y (other than 1 time for HW), blank - unclear/no response)
14) Did the student sketch from memory while studying? (0 - N, 0.5 - in head, 1 - Y, blank - unclear/no response)
15) Did the student gesture during description of their answer? (0 - N, 1 - Y, blank - unclear/no response)
16) Did the student use spatial reference words? (0 - N, 1 - Y, blank - unclear/no response)
    a) **Be careful they are not just terms that are part of a definition****

17) Explicit Mental image (0 - N, 1 - Y, blank - unclear/no response

   a) As defined by Emily: “a visual representation a student recalls and can use when solving a problem; a student is able to extract only the information that is immediately available from the surface level of the image.”

18) Is this image a sketch? (0 - N, 1 - Y, blank - unclear/no response)

19) Did the student describe insertion of question specific information to a mental or drawn image? (0 - N, 1 - Y, blank - unclear/no response)

   a.) In other words, did they use a previously known sketch straight from memory or **did they alter it**s appearance to help them with the problem (either in their head or on paper)

20) Did the student draw during problem solving? (0 - N, 1 - Y, blank - unclear/no response)

   a.) Specifically they begin to draw prior to obtaining the full solution, and use the image to help them? Including for confirmation

21) During the interview, does the student consult the sketch as they solve or explain, or use it to correct themselves? (0 - N, 1 - Y, blank - unclear/no response)

   a.) Be careful here...they might correct themselves, but sometimes it is just because they misread their own response, or it is unclear if the sketch told them or if something else did

22) Explicit mental model (0 - N, 1 - Y, blank - unclear/no response)

   a.) Mental model as defined by Emily = “a visual representation a student recalls that functions as an organizing structure for the student’s knowledge of a system. A highly developed visual mental model integrates information about structures, functions, and spatio-temporal processes, allowing the student to solve a problem or predict hypothesis by manipulating (modeling) the visual representation.”

A3. Code for Analyses completed in R

```
#Load requried packages
autopack <- function(pkg){
  new.pkg <- pkg[!(pkg %in% installed.packages()[, "Package"])]
  if(length(new.pkg))
    install.packages(new.pkg, dependencies = TRUE)
  sapply(pkg, require, character.only = TRUE)
}

packageslist <- c("tidyr", "dplyr", "knitr", "ggplot2", "purrr", "mice"
, "miceadds", "psy", "MASS")

autopack(packageslist)

#Import data file of all partially cleaned survey responses & fill blan
ks with NA
survey_responses <- read.csv("C:/Users/Jessica/Desktop/R/WM Thesis/Thes
is_data_R.csv", header = TRUE)
survey_responses <- read.csv("C:/Users/Jessica/Desktop/R/WM Thesis/Thes
is_data_R.csv", header = TRUE, nrows = sum(!is.na(survey_responses$Fini
shed)), na.strings=c("", "NA"))

#add columns where Likert scale items are converted to #s (and reversed
if needed), so they can be used for analysis
mapping7 <- c("Strongly disagree" = 1, "Disagree" = 2, "Somewhat disagr
ee" = 3, "Neither agree nor disagree" = 4, "Somewhat agree" = 5, "Agree
" = 6, "Strongly agree" = 7, "NA" = NA)

mapping5 <- c("Not confident at all" = 1, "Slightly confident" = 2, "So
mewhat confident" = 3, "Fairly confident" = 4, "Extremely confident" =
5, "NA" = NA)

#relevel 7-pt Likert scale for C, I, and U Qs
relvl_Likert7 <- function(x) {
x = factor(x,levels(x)[c(7,2,5,3,4,1,6)])
return(x)
}
survey_responses[53:61] <- lapply(survey_responses[53:61], relvl_Likert
7)

#relevel 5-pt Likert scale for SE Qs
relvl_Likert5 <- function(x) {
{x = factor(x,levels(x)[c(3,4,5,2,1)])
return(x)}
}
survey_responses[62:66] <- lapply(survey_responses[62:66], relvl_Likert
5)
```

```
#Convert 7-pt & 5-pt Likert to numerical
survey_responses[83:91] <- lapply(survey_responses[53:61], function(x){
mapping7[x]})
survey_responses[92:96] <- lapply(survey_responses[62:66], function(x){
mapping5[x]})

#reverse motivation Likert scale for 3rd Cost question so it is compara
ble to others
survey_responses$C3.1 <- survey_responses$C3.1*-1+8

#Make average columnsfor each construct
survey_responses$Cost <- rowMeans(survey_responses[,c("C1.1","C2.1","C3
.1")])
survey_responses$Interest <- rowMeans(survey_responses[,c("I1.1","I2.1"
,"I3.1")])
survey_responses$Utility <- rowMeans(survey_responses[,c("U1.1","U2.1",
"U3.1")])
survey_responses$SelfEff <- rowMeans(survey_responses[,c("SE1.1","SE2.1
","SE3.1", "SE4.1", "SE5.1")])

#view table to make sure it read and updated correctly
View(survey_responses)
```

Subset survey into batches

```
#2018 cohort (drawing intervention)
pre_survey_18 <- dplyr::select(filter(survey_responses, Timepoint == "P
re_18"), c(1:ncol(survey_responses)))
post_survey_18 <- dplyr::select(filter(survey_responses, Timepoint == "
Post_18"), c(1:ncol(survey_responses)))
One_sem_18 <- dplyr::select(filter(survey_responses, Timepoint == "One_
sem_18"), c(1:ncol(survey_responses)))
Cohort18_only <- rbind(pre_survey_18, post_survey_18, One_sem_18) #all
2018 cohort

#2019 cohort (comparison)
pre_survey_19 <- dplyr::select(filter(survey_responses, Timepoint == "P
re_19"), c(1:ncol(survey_responses)))
post_survey_19 <- dplyr::select(filter(survey_responses, Timepoint == "
Post_19"), c(1:ncol(survey_responses)))
One_sem_19 <- dplyr::select(filter(survey_responses, Timepoint == "One_
sem_19"), c(1:ncol(survey_responses)))
Cohort19_only <- rbind(pre_survey_19, post_survey_19, One_sem_19) #all
2019 cohort

#Subsetted motivation: only includes All_Draw and all numerical motivat
ion scores
Motiv2018 <- subset(One_sem_18, select = c(32,83:100))
```

```
Motiv2019 <- subset(One_sem_19, select = c(32,83:100))
MotivOnly <- rbind(Motiv2018, Motiv2019)
```

Prep for multiple imputations

```
#NOTES: m is the # of imputation -- default = 5; seed set just to make
sure it does the same each time; default for numerical is pmm (predicti
ve mean matching)

#View missing data patterns
md.pattern(MotivOnly)
```

Run multiple imputations on motivation and associated % study time (All_draw)

```
imp0 <- mice(MotivOnly, m=5, maxit = 0, seed = 123)
predMatrix <- imp0$predictorMatrix
#Exclude Mean construct scores as imputed variable
predMatrix["Interest",] <- 0
predMatrix["Cost",] <- 0
predMatrix["Utility",] <- 0
predMatrix["SelfEff",] <- 0
#Exclude Mean construct scores as predictor variable
predMatrix[,"Interest"] <- 0
predMatrix[,"Cost"] <- 0
predMatrix[,"Utility"] <- 0
predMatrix[,"SelfEff"] <- 0

#apply mult. imp. to without the mean construct score columns
impMiss <- mice(MotivOnly, m = 5, maxit = 10, pred = predMatrix, printF
lag = FALSE, seed = 123)
#Show info on which imputation methods used
print(impMiss)

#convert to dataframe to recalculate the mean columns, then turn back i
nto .mids
imp1 <- data.frame(complete(impMiss, include = TRUE, action = "long"))
imp1$Cost <- rowMeans(imp1[,c("C1.1","C2.1","C3.1")])
imp1$Interest <- rowMeans(imp1[,c("I1.1","I2.1","I3.1")])
imp1$Utility <- rowMeans(imp1[,c("U1.1","U2.1","U3.1")])
imp1$SelfEff <- rowMeans(imp1[,c("SE1.1","SE2.1","SE3.1", "SE4.1", "SE5
.1")])
imp2 <- as.mids(imp1)
```

Multiple regression of Motivation on drawing

```
#run the actual pooled models for the mult. regression; (lm(y ~ x1 + x2
+ x3, data=mydata)
MotivRegr <- with(imp2, lm(All_draw ~ Utility + SelfEff + Interest + Co
st))
```

```r
#show results
pool(MotivRegr)
summary(pool(MotivRegr))
pool.r.squared(MotivRegr, adjusted = FALSE)

#pearson inference correlation for multiply imputed datasets (on the 4
x variables: cost, utility, interest, self efficacy)
corr_MotivRegr <- miceadds::micombine.cor(mi.res = imp2, variables = c(
1,16:19))
corr_MotivRegr

#means of each
mUtil <- with(imp2, mean(Utility))
mSE <- with(imp2, mean(SelfEff))
mInt <- with(imp2, mean(Interest))
mCost <- with(imp2, mean(Cost))
mDraw <- with(imp2, mean(All_draw))

mean(unlist(mUtil$analyses))
mean(unlist(mSE$analyses))
mean(unlist(mInt$analyses))
mean(unlist(mCost$analyses))
mean(unlist(mDraw$analyses))

#SD of each
SDUtil <- with(imp2, sd(Utility))
SDSE <- with(imp2, sd(SelfEff))
SdInt <- with(imp2, sd(Interest))
SdCost <- with(imp2, sd(Cost))
SdDraw <- with(imp2, sd(All_draw))

mean(unlist(SDUtil$analyses))
mean(unlist(SDSE$analyses))
mean(unlist(SdInt$analyses))
mean(unlist(SdCost$analyses))
mean(unlist(SdDraw$analyses))

# mock plot this because it is actually 5 separate regressions pooled t
ogether; I used the first interation for this (and the others below), b
ut it is barely different for any others
par(mfrow = c(2,2))
plot(lm(All_draw ~ Utility + SelfEff + Interest + Cost, data=imp1[138:2
74,]))

#visualize general patterns in each construct (see above comment)
plot(All_draw ~ Utility, data=imp1[138:274,])
plot(All_draw ~ SelfEff, data=imp1[138:274,])
plot(All_draw ~ Interest, data=imp1[138:274,])
plot(All_draw ~ Cost, data=imp1[138:274,])
```

```r
#plot the simple linear regression of only cost
par(mfrow = c(1,1))

ggplot(imp1[138:274,], aes(x=Cost, y=All_draw)) +
  geom_point(color='black', size = 2) +
  geom_smooth(method=lm, color='red') +
  labs(y="Percent study time using Drawing", x="Cost") +
  theme(axis.title.x = element_text(size =15, vjust = 0), axis.title.y
= element_text(size =15, vjust = 1), axis.text = element_text(size =12,
color = "black")) +
  scale_y_continuous(breaks = seq(0,70,10)) + scale_x_continuous(breaks
= seq(1,7,1))

#simple linear regression model on only cost
justCost <- with(imp2, lm(All_draw ~ Cost))
pool(justCost)
summary(pool(justCost))
pool.r.squared(justCost, adjusted = FALSE)
```

Cronbach's alpha for one semester post motivation

```r
#before multiple imputations used to complete data set (all questions w
ithin a construct)

#for 2018 cohort (drawing)
cronbach(Motiv2018[c(2,4,7)]) #interest
cronbach(Motiv2018[c(5,8,9)]) #utility
cronbach(Motiv2018[c(3,6,10)]) #cost
cronbach(Motiv2018[c(11:15)]) #self efficacy

#for 2019 cohort (non-drawing)
cronbach(Motiv2019[c(2,4,7)]) #interest
cronbach(Motiv2019[c(5,8,9)]) #utility
cronbach(Motiv2019[c(3,6,10)]) #cost
cronbach(Motiv2019[c(11:15)]) #self efficacy

#for combined
cronbach(MotivOnly[c(2,4,7)]) #interest
cronbach(MotivOnly[c(5,8,9)]) #utility
cronbach(MotivOnly[c(3,6,10)]) #cost
cronbach(MotivOnly[c(11:15)]) #self efficacy
```

Simple linear model for drawing vs course grade

```r
#NOTES: this is not with mult. imp. because there is only a subset that
we collected data on grades, therefore it is not missing at random, not
necessary

Post18_grades <- read.csv("C:/Users/Jessica/Desktop/R/WM Thesis/Post18_
grades.csv", header = TRUE)
```

```r
Post18_grades <- read.csv("C:/Users/Jessica/Desktop/R/WM Thesis/Post18_
grades.csv", header = TRUE, nrows = sum(!is.na(Post18_grades$Progress))
, na.strings=c("", "NA"))

#plot
ggplot(Post18_grades, aes(x=All_draw, y=Grade_Percent)) +
  geom_point(color='black', size = 2) +
  geom_smooth(method=lm, color='red') +
  labs(y="Final course point total (%)", x="Percent study time using Dr
awing") +
  theme(axis.title.x = element_text(size =15, vjust = 0), axis.title.y
= element_text(size =15, vjust = 1), axis.text = element_text(size =12,
color = "black")) +
  scale_y_continuous(breaks = seq(40,100,10)) + scale_x_continuous(brea
ks = seq(0,100,10))

# build linear regression model on full data
coursePerform <- lm(Post18_grades$Grade_Percent ~ Post18_grades$All_dra
w)
print(coursePerform)
summary(coursePerform)
```