2022

Picture This! Arts Integration in the Classroom to Aid Retention of Abstract Science Concepts

Chloe Gorman

Picture This! Arts Integration in the Classroom to Aid Retention of Abstract Science Concepts

Chloe Gorman

Scripps College

November 23, 2022

Advised by Professor Johnson

& Professor Kovitz

Submitted to the Linguistics and Cognitive Science Department at Pomona College in partial fulfillment of the Bachelor of Arts Degree

Abstract

Previous cognitive science research has shown the "drawing effect," in which people tend to have better recall for something they've drawn rather than written. The drawing effect is an example of a "desirable difficulty," as drawing often takes more mental effort than reading or writing. This study explored whether the drawing effect would persist for more abstract words, which might be overly difficult to draw; rather, they might be a difficulty that is no longer desirable and instead just a source of distraction or cognitive overload. Participants were presented with more abstract and less abstract science terms and alternated drawing and writing each word. Their recall was tested in two memory tasks, 24 hours apart. Contrary to what was hypothesized, there was no drawing effect shown in this study. However, there was a significant difference in recall for words that were "more abstract" versus "less abstract," as well as a significant difference in words recalled in the second vocabulary test versus the first vocabulary test. Still, previous studies have shown great benefits from drawing and arts-integration for learning; future studies should explore the most effective ways to incorporate art and drawing into various educational settings.

keywords: drawing effect, desirable difficulty, learning, study method, arts-integration

Introduction

Much of the cognitive science research being done in the context of the classroom centers around memory and retention of the material being learned. There are various learning strategies that have been studied and recommended to improve learning and retention, such as spacing out learning sessions, interleaving different subjects and types of material, and doing self-quizzing and retrieval practice. There have been fewer studies that focus on integrating art practices, such as dancing, creative writing, and illustrating, into the classroom and other learning environments. The studies about arts-integration that *have* been done so far show that art can be a helpful tool for creating a deeper understanding of classroom subjects such as literature and science. Drawing, or illustrating concepts, has been shown to be an effective learning strategy both in and out of the classroom for long term retention of information (Fernandes et al, 2018). This is called the "drawing effect." When individuals illustrate a concept or piece of information, they tend to be able to recall it better than when they just hear or write out the concept in words. Drawing, instead of just writing, while taking notes in the classroom might be an optimal learning technique.

That being said, there is not as much research on the limitations of the drawing effect – especially when it comes to the type of stimuli. For example, do older students benefit from the drawing effect in the classroom? Can the drawing effect be utilized by students who are learning more abstract concepts, such as scientific theories of time or energy, that might be harder to illustrate and therefore too difficult and distracting? Alternatively, would the drawing effect be even more useful for learning concepts like these that are more abstract?

General Memory Strategies and Desirable Difficulties

There tends to be a general misconception about what "learning" truly is. Many students and educators view learning as the ability to regurgitate information on an assessment. This often leads to cramming for an assessment, not learning the material deeply but just remembering it long enough to take a test on it, and then immediately forgetting the information after the test. Brown, Roediger, and McDaniel, the authors of *Make It Stick* (2014), define learning as something that "requires memory," is an "acquired skill," and is "deeper when effortful" even though these can often feel unintuitive to students in a classroom setting.

One popular study method used by students is rereading material. Because rereading involves revisiting content that was already learned, it feels easy and familiar. In turn, students might have a false sense of their understanding of the material because the familiarity of rereading feels like remembering. Callender and McDaniel (2008) explain why rereading is not a beneficial way to learn and retain information. In their study, participants either read an educational passage once through or did "massed rereading" in which they read the passage twice without taking notes or using any other study techniques. The participants completed multiple-choice and short-answer assessments as well as summarized the material they read in order for the researchers to measure recall of the information. The researchers found that participants gained little to no benefit after rereading the material, no matter their individual reading comprehension.

Brown, Roediger, and McDaniel (2014) introduce the idea of "desirable difficulties," which reinforce different aspects of the material or skills being learned through effort. The authors explain, "effortful recall of learning…requires that you "reload" or reconstruct the components of the skill or material anew from long-term memory rather than mindlessly repeating them from short-term memory" (pp. 82). Desirable difficulties aid learning and retention because consolidating and reconsolidating information can only occur with effort. The authors suggest that failure can be a useful learning tool because it allows students to reconsolidate the information they've studied before, which strengthens the connections they have between what they already retained and what is still to be learned or understood.

One strategy that Brown, Roediger, and McDaniel (2014) suggest is spacing out study sessions as a desirable difficulty. Many students only do "blocked" practice, in which they do all of the studying at once, which has been shown in many studies to only be beneficial for more immediate recall. Blocked practice takes less cognitive effort (because there is less forgetting involved, as occurs between study sessions) and learners are often able to recall more immediately than learners who did spaced practice. However, spaced practice has been shown to provide significantly better long-term retention and recall than blocked practice. Because of this, spaced practice might seem unintuitive to learners since it feels more difficult and leads to worse recall than blocked practice at first. This is what makes spaced practice a desirable difficulty, as it demands more cognitive effort (through forgetting) but eventually allows for deeper learning and retention of information.

One study conducted by Bloom and Shuell (1981) shows the benefits of spaced practice through learning French vocabulary words. Participants (who were not native speakers of French) studied vocabulary words in either one massed block of practice for 30 minutes or three spaced blocks of 10 minutes each on three separate days. Bloom and Shuell (1981) found that participants in the blocked study group had better immediate recall than those in the spaced group; but, participants in the spaced study group had significantly better long-term retention and retrieval than those in the blocked study group when tested again four days later.

Soderstrom and Bjork (2015) explain that desirable difficulties can feel counterintuitive and inefficient. Many learners judge their learning and understanding based on their performance while they're learning and recalling. Soderstrom and Bjork (2015) assert that performance is not a great indicator of deep learning, especially when desirable difficulties are being utilized. They reference a study in which participants learned various movements to knock down a barrier on a hinge (Shea & Morgan, 1979). The blocked practice group practiced each of the three movement patterns separately, or in blocks. The interleaved practice group practiced each of the three movements patterns randomly mixed, rather than in their separate blocks. The participants in the blocked practice group performed better than those in the interleaved group after the first day. But, when the participants came back after 10 days to test their skills again, those in the interleaved group had a significantly better performance than those in the blocked group. This shows that although blocked practice might feel the best at first because it's easier, interleaving leads to improved long-term learning (even of a physical skill, rather than remembering verbal information). Interleaved practice might feel more difficult and inefficient (and actually produces worse results at first), but has been shown to be beneficial in the long-term, which is usually the goal of learning.

Bjork and Bjork (2011) also explain this difference by differentiating between performance versus learning and relating the two to storage strength versus retrieval strength. Performance is a measurement of more temporary "learning" or ability; in the case of the Soderstrom and Bjork (2015) study, the participants' ability directly after learning the movement patterns. This is closer to retrieval strength, which fluctuates and only represents the *current* level of accessibility of a memory or skill. Storage strength, on the other hand, is more permanent and shows how deeply something has been learned. Performance demonstrates retrieval strength, while actual learning and long-term recall ability is more a product of storage strength.

Elaboration

Especially relevant to arts-integration, elaboration is another desirable difficulty that has been found to aid long-term retention. Weinstein et al (2019) define elaboration as "add[ing] something to a memory" (102). Adding information to an existing memory allows for the content to be processed in a deeper, more permanent way. One type of elaboration is called "dual-coding," which involves studying material in multiple modes (i.e. visual *and* verbal) rather than just one mode. This is a form of elaboration because the information is processed in multiple ways, and can therefore be remembered later in multiple ways, allowing for easier access and deeper consolidation of the material.

Weinstein et al (2019) cite a study by Mayer and Anderson (1992), in which they found the benefits of dual-coding through animations paired with a narration of the animation (either of a brake system or a bicycle tire pump). The researchers found that participants had better recall for the information when they were presented with both a visual animation and an oral narration simultaneously than studying by only looking at an animation or only listening to a narration.

Dunlosky et al (2013) provide an overview of much of the research that has been done surrounding desirable difficulties. They cover ten learning techniques (some of which are desirable difficulties): elaborative interrogation, self-explanation, summarization, highlighting, mnemonics, imagery, rereading, testing, spacing, and interleaving. Based on Dunlosky et al's (2013) review of the literature at the time the article was written (published in 2013), Table 1 shows that desirable difficulties (elaborative interrogation, self-explanation, summarization, highlighting, the keyword mnemonic, imagery use for text learning, practice testing, distributed practice, and interleaved practice) are more useful for learning and retention than learning techniques that don't involve desirable difficulties (rereading):

Table 1. Study techniques rated by benefit for long-term retention

Technique	Utility
Elaborative interrogation	**Moderate**
Self-explanation	**Moderate**
Summarization	**Low**
Highlighting	**Low**
The keyword mnemonic	**Low**
Imagery use for text learning	**Low**
Rereading	**Low**
Practice testing	**High**
Distributed practice	**High**
Interleaved practice	**Moderate**

Note: This table was adapted from Table 4 (Utility Assessment and Rating of Generalizability for Each of the Learning Techniques) from the Dunlosky et al (2013) study.

Different learning strategies might be at different levels of being a desirable difficulty. It's possible that the learning techniques with the highest rated utility (practice testing and distributed practice) involved a desirable difficulty that took more effort than the other desirable difficulties (but weren't *too* difficult and didn't cause distraction or cognitive overload). Elaborative interrogation, self-explanation, and interleaving were rated "Moderate" for utility, and also involve more of a desirable difficulty than the rest of the techniques that were rated "Low," such as summarization and rereading.

The Picture-Superiority Effect

Paivio, Rogers, and Smythe (1968) introduce the "picture-superiority effect," which is that individuals tend to remember information better when it's presented in the form of an image rather than verbally or written in words. Participants had significantly better recall in a later free-recall test for the images than the words that were labels of the images (i.e. an image of a lobster versus the word "lobster"). The authors discuss the role of dual-coding in these findings, and that it might be the case that participants who were shown images were able to recall the information both visually and verbally since they already process a word to match each image automatically. By viewing an image, the individual perceives both the visual aspects of the image as well as the word they automatically pair with the image (i.e. when they see an image of a lobster, they automatically think "lobster").

The Drawing Effect

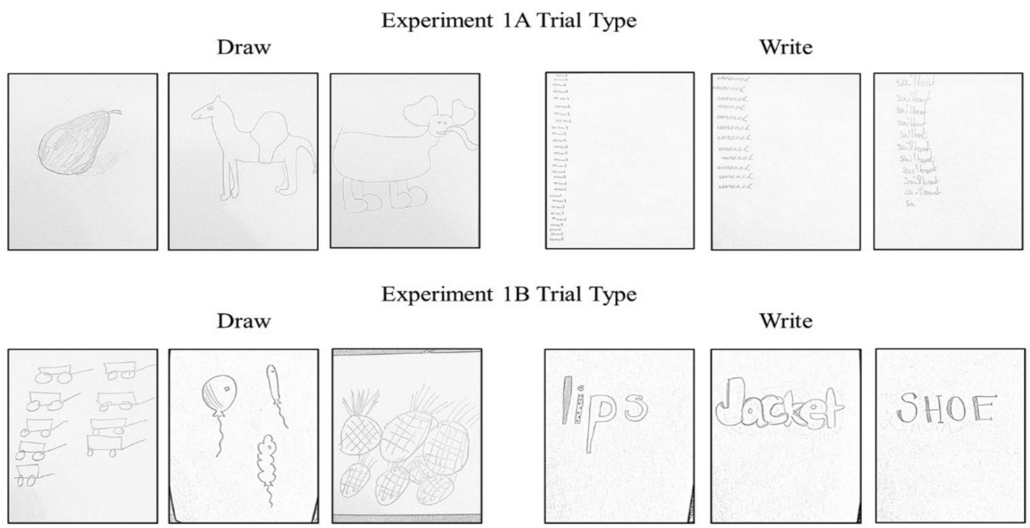
In terms of elaboration and dual-coding, a step above the picture-superiority effect is the "drawing effect" (Fernandes et al 2018). The drawing effect suggests that individuals tend to remember information better when they physically draw or illustrate it, rather than just viewing it as an image or reading it as a word. Drawing has three major beneficial components for memory:

pictorial, motoric, and elaborative. The pictorial factor refers to drawings being images or pictures; the viewer benefits from the picture-superiority effect just from this single factor of a drawing. Adding onto that, the motoric factor refers to the physicality of creating a drawing. This is a form of dual-coding, as the person drawing is both encoding the information both verbally in their mind as well as physically with their hand, giving them multiple paths to retrieve the information later on. The elaborative aspect of the drawing effect comes from translating a concept or word into an image. This requires the person drawing to think deeply enough about the material to represent it in a different form.

These three factors are helpful on their own, but are even more beneficial for long-term retention when they occur simultaneously; drawing allows this combination to occur. In their study, Wammes et al (2016) compare multiple combinations of these three factors: drawing without looking at the drawing (the intersection of motoric and elaborative), tracing a picture (the intersection of motoric and pictorial), and imagining a drawing without physically drawing it (the intersection of elaborative and pictorial). Drawing is unique because it is the intersection of all three; this was found to be the most beneficial for long-term retention of information, above any other combination of the three.

Wammes et al (2016) also found the drawing effect to be especially impactful for learning, above other variations of desirable difficulties: writing the stimuli repeatedly (Experiment 1A), writing out the stimuli in more detailed lettering (Experiment 1B), listing semantically similar words to allow for elaboration (Experiment 3), visualizing imagery instead of drawing (Experiment 4), and viewing images instead of drawing them (Experiment 5). Figure 1 shows what the participants' drawings and writing looked like. Words were either drawn or written repeatedly in Experiment 1A; in Experiment 1B, words were either repeatedly drawn or written with “added detail” rather than repeated:

Figure 1. Examples of participants’ draw and write trials (experiments 1A and 1B)



Note: This figure was provided by Wammes et al (2016) in Appendix B.

In all conditions, the words that were drawn were better recalled than the words that were written. In Experiment 6, it was also found that the drawing effect persists when the encoding time of stimuli is shorter (four seconds rather than forty). Wammes et al (2016) concluded that the drawing effect occurs because drawing “encourag[es] the cohesion of multiple modes of representation of target information…[to] create a more resilient trace, which is robust in the face of changes in setting, instructions, or competing encoding strategies" (pp. 1755). In other words, similar to Mayer and Anderson’s (1992) findings about dual-coding, the drawing effect provides a type of dual-coding by involving pictorial, motoric, and elaborative encoding. This also aligns with Fernandes et al (2018) findings that the drawing effect works because it combines these three factors simultaneously; isolating just one or two is not as effective for long-term retention.

The fact that the drawing effect is beneficial even under other learning limitations, such as shorter encoding times, might suggest that it could be beneficial under other challenges such as more complicated academic subjects or more abstract concepts that might be more difficult to illustrate.

General Benefits of Art for Memory and Aging Brains

Even beyond a classroom setting, arts-integration and the picture-superiority effect have been shown to be helpful for all age groups, even though current research about aging suggests that visual memory declines more quickly than verbal memory as a result of aging (Winograd et al 71, 1982). Winograd et al (1982) explain that if this is the case, the picture-superiority effect should be less beneficial for older participants with aging brains. Contrary to this, Winograd et al (1982) found that older participants had better recall for stimuli in the form of pictures rather than words, even though they had lower total recall in general than younger participants. The researchers concluded that the picture-superiority effect is still relevant to people with aging brains just as it benefits younger people.

Examples of Art Being Integrated into the Classroom

The Kennedy Center (2010) defines arts integration in the classroom as "an approach to teaching [in which] students construct and demonstrate understanding through an art form…students engage in a creative process which connects an art form and another subject area and meets evolving objectives in both" (pp. 1). Arts integration allows for learning to be more hands-on, collaborative, and creative, which all provide more opportunity for elaboration and deeper understanding of information.

Burnaford et al (2007) provide a review of the possible ways to integrate art into classrooms and other learning environments to deepen understanding and improve retention and engagement. Arts integration has been found to be especially helpful for improving literacy and reading comprehension. Burnaford et al (2007) mention the use of picture books to aid reading comprehension through connecting words and pictures, possibly as a form of dual-coding. Pictures in addition to words can increase engagement as well as deepen understanding when students are able to connect the text to the images. Andrzejczak et al (2005) conducted a study in which students drew images and then wrote about what they drew. The researchers found that the students' writing was more detailed and vivid after drawing and creating something first. This suggests that arts integration and the drawing effect could be beneficial for a wide range of subjects in school.

Crouch et al (2004) explore the benefits of simply changing the modes of presentation of information in the classroom setting. In their study, teachers either gave classroom physics demonstrations by having students just 1) watch the demonstration, 2) predict the outcome of the demonstration before watching, or 3) both predict the outcome *and* discuss their predictions with a partner before observing the demonstration and final outcome. The students rotated in groups through the conditions throughout the semester in order to be exposed to each teaching technique. They were assessed via a free-response test in order to measure how much they had retained from the demonstrations throughout the semester, based on the condition group they were in for each demonstration. Crouch et al (2004) found that students had the best recall for the demonstrations in which they were in the "discuss" condition. The next most beneficial condition was the "predict" group, with the "observe only" condition providing the lowest recall ability. The researchers suggest that these results imply significant benefits to increasing student engagement in the classroom. By predicting outcomes and discussing their thoughts with a partner, the students were able to utilize a form of elaboration through explaining their thought processes about the concepts they were learning, in addition to simply hearing the information verbally. Although the “predict” and “discuss” techniques aren’t directly related to art, they do imply that art could provide similar benefits through elaboration and increasing student engagement simply by varying the ways of presenting information in a science classroom.

Arts Integration and Science

Rosen O’Leary and Thompson (2019) explore the idea of including art in STEM (science, technology, engineering, and math) education as a way to benefit learning. They specifically suggest visual art, such as drawing. In their study, they found that participants who were visual note-takers (in other words, illustrated the science concepts they learned rather than just writing them out) performed better on a delayed recall test than the non-visual notetakers. Rosen O’Leary and Thompson (2019) emphasize the importance of arts-integration when it comes to learning STEM subjects specifically, as art has not been prioritized in the same way that STEM subjects are in school, nor have there been many studies conducted on the benefits of arts-integration in STEM classes.

Fello et al (2006) also explore the role of arts-integration in science classes, specifically in the form of “talking drawings.” In a 4th grade classroom, students were instructed to draw what they already knew about the material they were going to learn (about stars and constellations) and then discuss what they drew with a partner. After being given the lesson, they were told to draw again, this time incorporating what they had learned from the lesson. The teachers found that their second round of drawings allowed them to better notice what they didn’t understand before (such as the actual shape of a star, rather than the five-pointed star that one might think to draw before learning about stars in a scientific way). Students were then encouraged to compare their “pre-learning” and “post-learning” drawings to reflect on what they had learned and better understood through illustrating a concept twice. Fello et al (2006) also suggest that through drawing, the students were better able to visualize science concepts that had previously seemed too abstract. That being said, the "abstract" concepts they looked at were still being taught to students in elementary school: the ocean, stars, earth, etc. It would be interesting to take this a step further and study the effect of talking drawings when it comes to older students learning even more abstract science concepts that are difficult to visualize, such as energy or sound.

Hardiman et al (2019) look specifically at the role of arts integration for memory and recall of science content, rather than overall performance. Their study involved 16 4th and 5th grade classrooms. All students received an equal distribution of both arts and non-arts integrated learning, specifically for Astronomy, Life Science, Chemistry, and Environmental Science. Non arts-integrated learning involved teacher-led presentations, readings, and writing out notes rather than drawing and taking visual notes or singing about the material as in the arts-integrated condition. Similarly to previous studies, it was found that arts-integrated learning was more beneficial than the conventional non-arts integrated learning. Arts-integrated students were able to recall more information in the post-tests. Hardiman et al (2019) also found that students benefited the most when they were exposed to the arts-integrated learning condition first rather than second. They propose that this is the result of a "transfer effect" – that students learn how to learn through arts-integration and then use those skills in the non-arts integrated curriculum. Again, these studies were done with younger learners; a next step would be to study how the effects of arts-integration change when studying older students who are learning more complex or abstract science concepts.

The Present Study

In my study, I followed the methods of Experiment 1A of the Wammes et al (2018) study, with an added component of abstract science concepts as the stimuli being drawn or written. Participants were presented with the list of stimuli one word at a time and prompted to either repeatedly write out the word or illustrate the word. They were then assessed both five minutes later and 24 hours later (first via a free-recall test and then a vocabulary quiz each time). The independent variables were 1) if the word was drawn or written repeatedly by the participant, 2) if the word was in the "more abstract" list or the "less abstract" list, and 3) the timing of the two recall tasks (five minutes later versus 24 hours later). The two dependent variables were the recall frequency of the words in a free recall task, as well as frequency for correctly matching the words to their given definitions in a vocabulary test.

I specifically wanted to observe more abstract science concepts to understand the possible limitations to the drawing effect, or if it would still hold for concepts that are difficult to visualize or illustrate in hopes of illuminating the possibilities of integrating the drawing effect into a classroom environment in which students are learning more complicated and abstract concepts. For example, would the drawing effect be beneficial when it comes to note taking in a college physics class? Or, would drawing cease to be a desirable difficulty and become a distraction or source of cognitive overload in the classroom instead?

I hypothesized that the drawing effect would remain beneficial even when participants were studying more abstract concepts, both for recalling the words as well as their definitions. Based on the Wammes et al (2018) study, the drawing effect endured other types of interference such as having participants be in a lecture hall that would be more distracting than a quiet study room as in the other experiments (Experiment 2) or only allowing participants a much shorter encoding time (Experiment 6). This suggests that the drawing effect might not be negatively affected by the added difficulty of the words being more abstract and challenging for the participant to know how to represent in a drawing. The drawing effect in the Wammes et al (2018) article also remained beneficial even after being compared to other desirable difficulties such as listing semantically similar words to the word being studied (Experiment 3) and visualizing other images related to the words being studied (Experiment 4). Drawing was more beneficial in all eight of the experiments in the Wammes et al (2018) study *and* proved to be durable when added difficulties or distractions were involved. This would suggest that the added challenge of the words from the list of stimuli being more abstract wouldn't be too distracting or challenging to diminish the drawing effect.

In fact, this added difficulty of representing a more abstract concept could prove to be even *more* beneficial than for studying words that are less abstract and easier for participants to illustrate. As Brown, Roediger, and McDaniel (2014) concluded, desirable difficulties aid learning because learning requires effort; participants would be required to think more deeply about the more abstract concepts and make connections to the information they already know about the concepts in order to know how best to illustrate them. This form of elaboration is also supported by Fernandes et al (2018), who explain that the drawing effect is the result of three beneficial factors simultaneously: pictorial, motoric, and elaborative. Making the concepts even more abstract might add to the elaborative aspect of the drawing effect.

I also predicted that the words that were written would have the best recall only in the first recall task immediately after studying and that the words that were drawn would have the best recall in the second recall task, as this would be testing long-term retention and storage strength rather than just retrieval strength. As explained in the Bjork and Bjork (2011) article, as well as the Soderstrom and Bjork (2015) article, performance is not always a reliable indicator of actual learning, and desirable difficulties might feel like "worse" learning at first. Since the drawing effect would remain a desirable difficulty for participants studying more abstract words, I predicted a similar pattern in which there would tend to be a higher frequency of recall for the words written only at first; after 24 hours, however, I predicted that the words that were drawn would have a much better recall frequency, as participants would have a stronger storage strength of the drawn words that took more effort to encode. I expected this pattern to be the same for both the free recall and the vocabulary test.

Methodology

Participants

13 participants participated in this study. All participants were above age 18 and enrolled students at the Claremont Colleges. Participants were recruited through Facebook posts, a posting in the cognitive science listserv, and word of mouth. They were compensated $15 for their time.

Materials

Participants were presented with a video that showed fifteen "more abstract" and fifteen "less abstract" science terms and concepts, along with a short definition for each (see Appendix A for the list of stimuli). The words were pulled from three different study sheets for college-level classes: Biology for Dummies (Kratz 2022), AP Environmental Science (Schoolinsites.com), and College Chemistry (Wicks). I selected words that I thought most participants would be at least familiar with, as well as words that would be easy to label as either more abstract or less abstract. Participants were also given a blank worksheet with 30 boxes to draw and write the terms.

Procedure

Study task: Participants were shown the video with the list of stimuli, seeing one word and its definition at a time. Each term was shown for 40 seconds and they were instructed to read the word and its provided definition on the screen and were then prompted to either 1) repeatedly write each word or 2) illustrate the word for the full 40 seconds.

Filler task: Participants were administered a filler task, in the form of a nine-piece jigsaw puzzle online.

First retrieval task: After completing the filler task, they completed a free-recall test. They were instructed to type as many of the terms that they remembered seeing in the video. After the free-recall test, participants were given a vocabulary quiz to test their memory of the definitions in addition to the 30 words. They were given each definition and entered the term from the video that matched the definition.

Second retrieval task: Participants returned the next day (24 hours later) to complete another free-recall test with the same instructions to write the words they remembered from the list of stimuli. Participants then completed the same vocabulary quiz, matching the terms to the provided definitions.

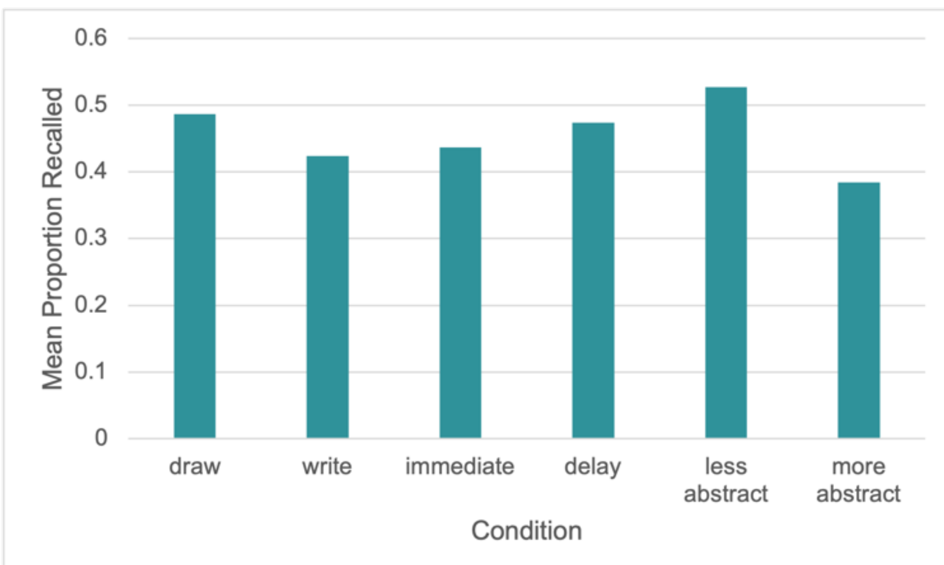
Outgoing survey: Participants completed a short survey asking about their usual study methods and their perceived level of difficulty, effectiveness, and enjoyment while using the study methods in this study. They also rated their level of familiarity from 1-7 with each word they studied from the video.

Results

For the free recall test, I used a 3-way ANOVA to analyze the effect of three independent variables on the number of terms that were recalled correctly: draw versus write, more abstract terms versus less abstract terms, and the first free recall test versus the second memory test 24 hours later.

As shown in Figure 2, for the free recall test, there was no significant difference in recall between the drawn words and the written words. In other words, there was no drawing effect shown in this study, $F(1, 12) = 2.348$, $p = .151$, $\eta_p^2 = 0.164$. There was also no significant difference between the number of words recalled in the first memory task compared to the second memory task 24 hours later, $F(1, 12) = 2.619$, $p = .132$, $\eta_p^2 = 0.179$. There was, however, a significant difference between the number of recalled terms that were more or less abstract, $F(1, 12) = 6.627$, $p = .024$, $\eta_p^2 = 0.356$. Words that were less abstract were recalled significantly more often than words that were more abstract.

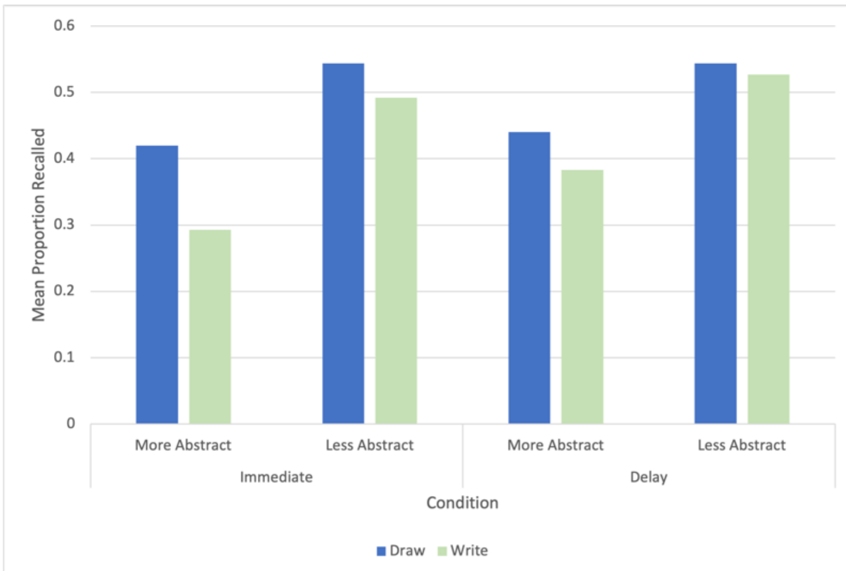
Figure 2. Marginal means for each variable in the free recall test



As shown in Figure 3, I used the 3-way ANOVA to analyze the interactions between the independent variables for the free recall test. There was no significant interaction between the drawing condition and delay, $F(1, 12) = 2.60$, $p = .133$, $\eta_p^2 = 0.178$. There was no significant interaction between the drawing condition and abstractness, $F(1, 12) = 0.917$, $p = .357$, $\eta_p^2 =$

0.071. There was also no significant interaction between the delay and abstractness, $F(1, 12) = 0.310$, $p = .588$, $\eta_p^2 = 0.025$. There was no significant 3-way interaction, $F(1, 12) = 0.161$, $p = .695$, $\eta_p^2 = 0.013$.

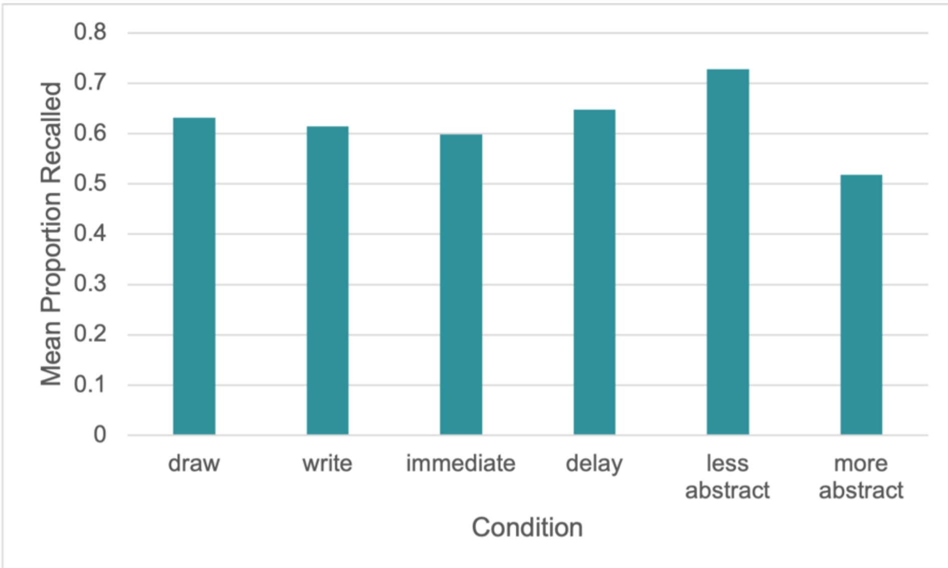
Figure 3. Means for each condition in the free recall test



I also used a 3-way ANOVA for the vocabulary tests to analyze the effect of three independent variables on the number of terms that were correctly matched to their definition: draw versus write, more abstract terms versus less abstract terms, and the first vocabulary test versus the second vocabulary test 24 hours later.

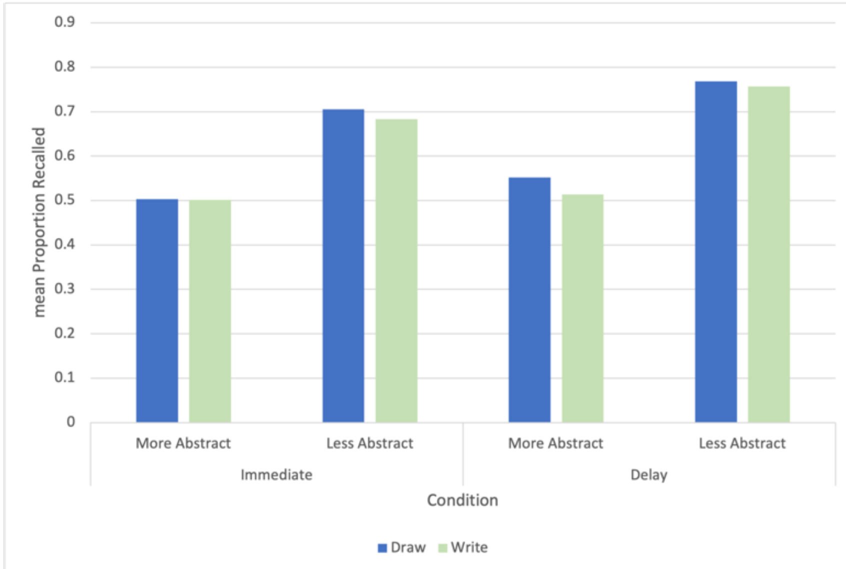
As shown in Figure 4, for the vocabulary test, there was no significant difference in recall between the words written versus the words drawn, $F(1, 12) = 0.179$, $p = .680$, $\eta_p^2 = 0.015$. Unlike in the free recall test, there was a significant difference in the number of words that were recalled in the first memory task compared to the second memory task 24 hours later, $F(1, 12) = 5.832$, $p = .033$, $\eta_p^2 = 0.327$. Participants recalled significantly more words in the second memory task than in the first. There was also a significant difference between the number of words recalled that were more abstract versus less abstract in the vocabulary test, $F(1, 12) = 31.250$, $p < .001$, $\eta_p^2 = 0.723$. Participants correctly recalled significantly more terms that were less abstract rather than more abstract.

Figure 4. Marginal means for each variable in the vocabulary test



As shown in Figure 5, in analyzing the interactions between the independent variables for the vocabulary test, there was no significant interaction between the drawing condition and delay, $F(1, 12) = 0.293$, $p = .598$, $\eta_p^2 = 0.024$. There was no significant interaction between the drawing condition and abstractness, $F(1, 12) = 0.001$, $p = .970$, $\eta_p^2 = 0.000$. There was no significant interaction between the delay and abstractness, $F(1, 12) = 2.415$, $p = .146$, $\eta_p^2 = 0.168$. Lastly, there was no significant 3-way interaction, $F(1, 12) = 0.759$, $p = .401$, $\eta_p^2 = 0.059$.

Figure 5. Means for each condition in the vocabulary test



I also analyzed the data from the survey questions about participants' enjoyment and perceived difficulty of drawing and writing the words during this study by using paired samples t-tests. There was a significant difference in the perceived difficulty of writing versus drawing, $t(12) = -4.62$, $p < .001$, $d = -1.282$. Participants rated perceived difficulty of writing and drawing from 1-7; the mean for difficulty writing was 2.31 ($SD = 1.49$) and the mean for difficulty drawing was 4.54 ($SD = 1.27$), so participants tended to rate drawing as more difficult than writing. There was also a significant difference in the reported enjoyment of writing versus drawing, $t(12) = -3.13$, $p = .009$, $d = -0.868$. The mean for reported enjoyment of writing was 3.08 ($SD = 1.55$) and the mean for reported enjoyment of drawing was 5.0 ($SD = 1.29$), so participants tended to rate drawing as more enjoyable than writing.

After completing the first memory task, participants made a prediction about their recall. 64.3% of participants said they thought their recall would be better for the words they drew, 21.4% predicted their recall would be the same for the words they drew and wrote, and 14.3% thought their recall would be better for the words they wrote.

Participants also rated their familiarity with each word on a scale of 1-7. Table 2 shows the average ratings of each word, which are organized in the table as “more abstract” or “less abstract.” The average familiarity ratings for more abstract versus less abstract words were similar, with the average familiarity rating for more abstract words being 4.419 out of 7 and the average familiarity rating for less abstract words being 4.668 out of 7. A paired samples t-test showed that there was no significant difference between the familiarity ratings for more abstract versus less abstract words, $t(15) = -0.757$, $p = .461$, $d = -0.196$.

Table 2. Participants’ familiarity ratings of each term.

Abstractness	Term	Average Rating (from 1-7)	Standard Deviation
More abstract	Homeostasis	5.07	2.018
More abstract	Evolution	6.71	0.504
More abstract	Natural selection	6.43	0.756
More abstract	Mutualism	4.86	1.791
More abstract	Biodiversity	5.71	1.204
More abstract	Species diversity	4.29	1.660
More abstract	Species evenness	2.71	1.489
More abstract	Rehabilitation	4	2.148
More abstract	Ecological tolerance	2.36	1.499
More abstract	Limiting resource	3.21	1.902
More abstract	Noise pollution	5.36	1.823
More abstract	Biomass	3.43	1.505
More abstract	Sustainable yield	2.79	1.626
More abstract	Law of Conservation of Mass	5.29	2.164
More abstract	Dalton’s Law	4.07	2.165
Less abstract	Heart	6.43	0.852
Less abstract	Skeleton	6.43	0.938
Less abstract	Mitosis	5.29	1.589
Less abstract	Central nervous system	5.5	1.225
Less abstract	Peripheral nervous system	4.5	1.829
Less abstract	Primary succession	2.43	1.555
Less abstract	Clearcutting	3.29	1.978

Less abstract	Waterlogging	2.5	1.787
Less abstract	Drip irrigation	3.14	2.107
Less abstract	Aquaculture	3.43	1.989
Less abstract	Coal	5.71	1.267
Less abstract	Crude oil	4.29	2.367
Less abstract	Geothermal energy	3.29	1.684
Less abstract	Predator	6.86	0.363
Less abstract	Prey	6.93	0.267

Discussion

I hypothesized that while immediate recall would be better for words that were written, delayed recall would be significantly better for words that were drawn. I also hypothesized that recall might be even better for words that were more abstract rather than less abstract, as the increased level of abstractness might add to the elaborative aspect of the drawing effect that has been found in many previous studies. Contrary to my hypothesis, I found no significant drawing effect, let alone recall being better for the more abstract words that were drawn than the less abstract words that were drawn. In fact, in both the free recall test and the vocabulary test, words that were less abstract were recalled significantly more often than words that were more abstract. Still, the words that were drawn were remembered slightly more than the words that were written, although the difference was not statistically significant.

This lack of a drawing effect might be a result of the word list that was used, rather than the durability of the drawing effect. Based on the free recall results, participants tended to remember certain words in the same groupings, possibly implying that some of the terms were cueing each other. For example, most participants remembered the word pair “predatory” and “prey” as well as the pair “central nervous system” and “peripheral nervous system." Because these word pairs were together in the main list of terms, they were then evenly split up between the draw and write conditions, as every other word in the list was assigned "draw" while every other word was assigned "write." This meant that whenever one of the terms in the pair was written, the other was always drawn and vice versa. If one of these terms was remembered, it often cued the other. Since there were only 30 terms, having any sort of memory cueing between words might have lessened the drawing effect; if there was a drawing effect, it might have been hidden by the words in the draw condition also cueing memory for the words in the write condition.

Interestingly, participants recalled significantly more words in the second vocabulary test than the first. This unexpected result might be the product of a "practice effect," as participants might have been reminded of words by taking the first vocabulary test as they were given the definitions and answered with the term. After then taking a second free recall test the next day, they would've been more familiar with the material and structure of the memory tasks. It's possible that this caused participants to then able to recall more in the second vocabulary test (and slightly more in the second free recall test, although this difference was not statistically significant).

The significant difference in the number of more abstract versus less abstract words that were recalled was also expected – participants remembered significantly more less abstract words in both the free recall and vocabulary tests. This might suggest that less abstract words are just easier to remember, as there was also no significant difference between the familiarity ratings for less abstract versus more abstract words. If this had been a significant difference, it might have been that the less abstract words were more familiar to participants, but this was not found.

Lastly, the ratings for the enjoyment and difficulty matched my hypothesis, as participants, on average, rated drawing to be more difficult but also more enjoyable than writing. Even though a significant drawing effect was not observed in this study, students might change their study strategies if they were aware of this potential drawing effect in their own studies. Few participants reported drawing as one of their main study methods and this might be due to the fact that drawing is perceived as more difficult than writing. If more students considered desirable difficulties in their learning, they might choose to incorporate drawing into the learning, as most participants also found it more enjoyable than writing. Interestingly, participants also predicted that they would better remember the words they drew over the words they wrote, even though there ended up being no significant drawing effect.

Limitations/Future Work

The most significant limitation to this study was the small sample size; the sample size of 13 participants was based on previous research having success in finding a strong drawing effect with similar numbers of participants, but it’s possible that having a larger number of participants might have been helpful in this study to find a significant drawing effect.

Similarly, it might be helpful to use words that are more difficult and novel to the participants. It’s possible that the terms were too easy and that participants were simply using their own knowledge to answer the vocabulary questions. For example, most participants in this study reported not being familiar with the term “Dalton’s Law” but every participant was familiar with the term “heart.” There might have been interference with participants’ previous knowledge of some of these terms; there might be a significant drawing effect if the terms had all been equally less familiar to participants, or even completely arbitrary or made-up words.

In order to avoid the potential practice effect that was seen by recall improving for both the delayed free recall task and the delayed vocabulary task, future studies could omit the vocabulary test from the memory task phase and have participants complete only an immediate and delayed free recall task. This way, the vocabulary test wouldn't remind participants of terms in the free recall test or give participants practice for the vocabulary test the next day.

Conclusion

Contrary to what was hypothesized, there was no drawing effect observed in this study; there was no significant difference in recall between drawn words and written words in either the free recall tests or the vocabulary tests. Still, there were slightly more drawn words recalled than written words; it's possible that this would've been a significant effect if there had been more participants and if the list of stimuli was written in a way that would avoid paired words cueing each other across the drawing and writing conditions. Although there was no significant drawing effect in this study, previous studies have shown strong, durable drawing effects. Additionally, the participants in this study reported feeling that drawing was more difficult but also more enjoyable. This matches what has been reported in previous studies and supports the idea that drawing is a desirable difficulty and can be utilized by students. Future studies should explore the most beneficial ways for drawing and other art forms to be integrated into classrooms and other learning environments.

References

Bjork, Elizabeth L. & Bjork, Robert A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). Worth Publishers.

Bloom, K. C., & Shuell, T. J. (1981). Effects of massed and distributed practice on the learning and retention of second-language vocabulary. *The Journal of Educational Research*, *74*(4), 245–248. https://doi.org/10.1080/00220671.1981.10885317

Brown, Peter C., Roediger, Henry L, & McDaniel, Mark A. (2014). *Make it stick: The science of successful learning.* Harvard University Press. http://www.jstor.org/stable/j.ctt6wprs3

Callender, A. A., & McDaniel, M. A. (2009). The limited benefits of rereading educational texts. *Contemporary Educational Psychology*, *34*(1), 30–41. https://doi.org/10.1016/j.cedpsych.2008.07.001

Crouch, C., Fagen, A., Callan, J., & Mazur, E. (2004). Classroom demonstrations: Learning tools or entertainment? *American Journal of Physics - AMER J PHYS*, *72*. https://doi.org/10.1119/1.1707018

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Fello, S. E., Raquette, K. R., & Jalongo, M. R. (2006). Talking drawings: Improving intermediate students' comprehension of expository science text. *Childhood Education*, *83*(2), 80–86. https://doi.org/10.1080/00094056.2007.10522885

Fernandes, M. A., Wammes, J. D., & Meade, M. E. (2018). The surprisingly powerful influence of drawing on memory. *Current Directions in Psychological Science*, *27*(5), 302–308. https://doi.org/10.1177/0963721418755385

Hardiman, M., JohnBull, R., Carran, D., & Shelton, A. (2019). The effects of arts-integrated instruction on memory for science content. *Trends in Neuroscience and Education*, *14*. https://doi.org/10.1016/j.tine.2019.02.002

Kratz, Rene F. (2022, Feb 25). *Biology for dummies cheat sheet.* Dummies.com. https://www.dummies.com/article/academics-the-arts/science/biology/biology-for-dummies-cheat-sheet-209035/#tab1

Mayer, R. E., & Anderson, R. B. (n.d.). *The instructive animation: Helping students build connections between words and pictures in multimedia learning*. 9.

Paivio, A., & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology*, *5*(2), 176–206. https://doi.org/10.1016/0010-0285(73)90032-7

Paivio, A., Rogers, T. B., & Smythe, P. C. (1968). Why are pictures easier to recall than words? *Psychonomic Science*, *11*(4), 137–138. https://doi.org/10.3758/BF03331011

Rosen-O'Leary, R., & Thompson, E. G. (2019). STEM to STEAM: Effect of visual art integration on long-term retention of science content. *Journal for Leadership and Instruction*, *18*(1), 32–35.

Schoolinsites.com Authors. n.d. *AP Environmental Science: 'Cheat' Sheet.* [PDF]. Schoolinsites.com. https://content.schoolinsites.com/api/documents/9f079bf42ef04ec3bfd41ff2b48fbc8f.pDf

Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 179–187.http://dx.doi.org/10.1037/0278-7393.5.2.179

Soderstrom, N. C., & Bjork, R. A. (2015). Learning versus performance: An integrative review. *Perspectives on Psychological Science*, *10*(2), 176–199. https://doi.org/10.1177/1745691615569000

Wammes, J. D., Meade, M. E., & Fernandes, M. A. (2016). The drawing effect: Evidence for reliable and robust memory benefits in free recall. *Quarterly Journal of Experimental Psychology*, *69*(9), 1752–1776. https://doi.org/10.1080/17470218.2015.1094494

Weinstein, Y., Sumeracki, M., & Caviglioli, O. (2019). Development of understanding. *Understanding how we learn: a visual guide* [PDF file]. (1st Ed., pp. 101-116). Routledge.

Wicks, Gene. n.d. *College Chemistry: College Chemistry First Semester Review Sheet.* [PDF]. Saultschools.org. https://www.saultschools.org/cms/lib/MI17000143/Centricity/Domain/162/College%20Chemistry%20First%20Semester%20Review%20Sheet.pdf

Winograd, E., Smith, A. D., & Simon, E. W. (1982). Aging and the picture superiority effect in recall. *Journal of Gerontology*, *37*(1), 70–75. https://doi.org/10.1093/geronj/37.1.70

Appendix A

<u>List of Stimuli:</u>

More abstract

1. Homeostasis: "the balance, or equilibrium, of the body" (Kratz)
2. Evolution: "the change of living things over time" (Kratz)
3. Natural Selection: "the theory that in any given generation, some individuals are more likely to survive and reproduce than others" (Kratz)
4. Mutualism: type of "relationship between two species in an ecosystem" in which "both benefit" (Schoolinsites.com)
5. Biodiversity: "variety of different species" (Schoolinsites.com)
6. Species Diversity: "number and variety of species in a given area" (Schoolinsites.com)
7. Species Evenness: "comparative number of individuals of each species present" (Schoolinsites.com)
8. Rehabilitation: "attempting to turn a degraded ecosystem back to being functional" (Schoolinsites.com)
9. Ecological Tolerance: "the range of conditions, such as temperature, salinity, flow rate, and sunlight that an organism can endure before injury or death results" (Schoolinsites.com)
10. Limiting Resource: "a resource that a population cannot live without and which occurs in quantities lower than the population would require to increase in size" (Schoolinsites.com)
11. Noise Pollution: "caused by unnecessary sounds that commonly travel through the air" (Schoolinsites.com)
12. Biomass: "energy derived from living organisms" (Schoolinsites.com)
13. Sustainable Yield: "the amount of a renewable resource that can be taken without reducing the available supply" (Schoolinsites.com)
14. Law of Conservation of Mass: "matter can be neither created nor destroyed" (Wicks)
15. Dalton's Law: "total pressure is equal to the sum of the partial pressures" (Wicks)

Less abstract

1. Heart: "a muscular pump. Contractions by the heart push blood through the circulatory system" (Kratz)
2. Skeleton: "provides support and structures for muscles to pull against" (Kratz)
3. Mitosis: "[a process that occurs when cells] are going to make an exact copy of themselves for asexual reproduction, growth, or tissue repair" (Kratz)
4. Central Nervous System: "made up of the brain and spinal cord and sends out instructions [to the rest of the body]" (Kratz)

5. Peripheral Nervous System: “contains the nerves that send the messages from the central nervous system to the rest of the body” (Kratz)
6. Primary Succession: “begins with a lifeless area where there is no soil (ex. bare rock). Soil formation begins with lichen or moss” (Schoolinsites.com)
7. Clearcutting: “removal of all trees from an area” (Schoolinsites.com)
8. Waterlogging: “excess water in soil, raises water table” (Schoolinsites.com)
9. Drip Irrigation: “water released through perforated hoses at plant roots” (Schoolinsites.com)
10. Aquaculture: “fish and aquatic plants farming” (Schoolinsites.com)
11. Coal: “a combustible rock consisting mainly of carbonized plant matter, found mainly in underground deposits and widely used as fuel” (Schoolinsites.com)
12. Crude Oil: “liquid petroleum pumped from the ground” (Schoolinsites.com)
13. Geothermal Energy: “is obtained by using the heat stored in the Earth’s interior to heat up water, which is brought back to the surface as steam” (Schoolinsites.com)
14. Predator: “organism that eats another organism” (Schoolinsites.com)
15. Prey: “organism eaten by predator” (Schoolinsites.com)

<u>Survey Questions</u>
What are your primary study/notetaking methods (i.e. writing, reading, drawing, reciting, self-testing, explaining, etc)?

How difficult would you rate repeatedly writing each word from 1-7?

How difficult would you rate drawing each word from 1-7?

How much did you enjoy writing each word that you wrote from 1-7?

How much did you enjoy drawing each word that you drew from 1-7?

Do you think your recall rate was higher for words that you wrote or drew?

Rate how familiar you are with each word from 1-7: